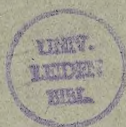


Robuste Schätzung der Varianz

Rainer Vollmerhaus

Disbui. München

1989





Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41554827_0086

Robuste Schätzung der Varianz

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

vorgelegt von

Rainer Vollmerhaus

1989

Referent:

Prof. Dr. Kurt Weichselberger

Korreferent:

Prof. Dr. Bernhard Rüger

Promotions-
abschlußberatung:

26.7.1989

Die Bedeutung des Verfalls

Die Bedeutung des Verfalls ist ein Thema, das in der Philosophie und in der Literatur seit Jahrhunderten diskutiert wird. In der Philosophie geht es darum, was das Leben bedeutet und wie wir es verstehen können. In der Literatur wird das Thema oft durch die Erzählung von Geschichten illustriert, die den Prozess des Verfalls zeigen.

Im Folgenden wird die Bedeutung des Verfalls in der Philosophie und in der Literatur diskutiert.

Die Bedeutung des Verfalls in der Philosophie

Die Bedeutung des Verfalls in der Literatur

Die Bedeutung des Verfalls in der Philosophie ist ein Thema, das in der Philosophie und in der Literatur seit Jahrhunderten diskutiert wird. In der Philosophie geht es darum, was das Leben bedeutet und wie wir es verstehen können. In der Literatur wird das Thema oft durch die Erzählung von Geschichten illustriert, die den Prozess des Verfalls zeigen.

Die Bedeutung des Verfalls in der Literatur

Die Bedeutung des Verfalls in der Literatur ist ein Thema, das in der Philosophie und in der Literatur seit Jahrhunderten diskutiert wird. In der Philosophie geht es darum, was das Leben bedeutet und wie wir es verstehen können. In der Literatur wird das Thema oft durch die Erzählung von Geschichten illustriert, die den Prozess des Verfalls zeigen.

Die Bedeutung des Verfalls in der Literatur

Die Bedeutung des Verfalls in der Literatur ist ein Thema, das in der Philosophie und in der Literatur seit Jahrhunderten diskutiert wird. In der Philosophie geht es darum, was das Leben bedeutet und wie wir es verstehen können. In der Literatur wird das Thema oft durch die Erzählung von Geschichten illustriert, die den Prozess des Verfalls zeigen.

Inhaltsverzeichnis

1. Einleitung	1
2. Grundlagen	4
2.1. Die Gammaverteilung	4
2.2. Schätzer im Gammaverteilungsmodell	9
2.3. Die Verteilung von Q	11
2.4. Die Verteilung von $\hat{a}_{ML}$	17
2.5. Die Verteilung von $\ln Q$	21
2.6. Die asymptotische Verteilung von $\hat{a}_{ML}$	25
2.7. Die asymptotische Verteilung von $\hat{\sigma}_{ML}^2$	27
2.8. Einige Eigenschaften der Verteilungsfunktion von $\hat{a}_{ML}$	29
2.9. Einige Eigenschaften der Verteilungsfunktion $\mathcal{G}_{a,\sigma^2}$	30
3. Die Berechnung der Verteilungsfunktion von Q	32
3.1. Die Berechnung mit Hilfe orthogonaler Polynome	32
3.1.1. Die allgemeine Methodik	33
3.1.2. Die Entwicklung nach verallgemeinerten Laguerre-Polynomen	35
3.1.3. Die Entwicklung nach verallgemeinerten Jacobi-Polynomen	40
3.2. Die Berechnung mit Hilfe einer Laplace-Transformation	43
3.3. Einige numerische Beispiele	45
4. Ein einseitiges Konfidenzintervall für die Varianz	50
4.1. Modellbeschreibung und Zielvorstellung	50
4.2. Die Form des Konfidenzintervalls	52
4.3. Das Konstruktionsprinzip	58
4.4. Die indirekte Bestimmung von $C(\underline{X})$	61
4.5. Ein erster Versuch	64
4.6. Ein bedingtes Konfidenzintervall	66
4.7. Die Wahl der Startfunktion	67
4.8. Eigenschaften der Startfunktion $c^{(1)}$	81
4.9. Der Iterationsschritt von $c^{(i)}$ nach $c^{(i+1)}$	86
4.10. Einige Beispiele	93
4.11. Überlegungen zu Verallgemeinerungsmöglichkeiten	99
5. Ein Verfahren zur Abschätzung der Überdeckungswahrscheinlichkeiten	102
5.1. Die Abschätzung der Verteilungsfunktion von $\bar{X}$	103
5.2. Die Abschätzung der Verteilungsfunktion von Q^a	107
5.3. Die Abschätzung der Überdeckungswahrscheinlichkeiten	110
5.4. Einige Beispiele	114
6. Zusammenfassung	120
Symbolverzeichnis	121
Literaturverzeichnis	122

1. Einleitung

Motivation

Die vorliegende Arbeit wurde durch die im folgenden beschriebene Fragestellung angeregt. Für den Erwartungswert einer rechtsschiefen Verteilung soll ein zweiseitiges Konfidenzintervall konstruiert werden, wobei die Varianz der zugrundeliegenden Verteilung unbekannt sei. Der Umfang der zur Bildung des Konfidenzintervalls heranzuziehenden Stichprobe soll so gewählt werden, daß eine vorgegebene Länge bzw. ein vorgegebener Erwartungswert für die Länge des Konfidenzintervalls nicht überschritten wird. Diese Länge bzw. deren Erwartungswert wird in vielen Fällen — wie auch in der Normalverteilungssituation — zumindest näherungsweise proportional zum Quotienten aus Standardabweichung und der Wurzel des Stichprobenumfangs sein. Es ist also in einer solchen Situation sinnvoll, zur Planung des Stichprobenumfangs die Varianz der Grundgesamtheit aus einer Vorstichprobe abzuschätzen und zwar — um auf der sicheren Seite zu sein — nach oben. Eine solche aus der Vorstichprobe bestimmte obere Schranke kann dabei natürlich auch einmal unter dem der Verteilung der Grundgesamtheit zugrundeliegenden, wahren Wert der Varianz liegen. Um solche „Fehlschätzungen“ zu kontrollieren, wird die obere Schranke im Rahmen der Bestimmung eines einseitigen Konfidenzintervalls zu einem vorgegebenen Konfidenzniveau ermittelt. Solch ein Konfidenzintervall für die Varianz soll in dieser Arbeit konstruiert werden. Da der Umfang der Vorstichprobe sinnvollerweise nicht zu groß sein sollte, scheiden dabei asymptotische Methoden zur Lösung des Problems aus.

Es wird davon ausgegangen, daß die zugrundeliegende Verteilung eine Gammaverteilung ist, wobei beide Parameter der Verteilung — der Gestaltparameter und der Skalierungsparameter — als unbekannt vorausgesetzt werden. Die Wahl dieses Modells begründet sich durch die hohe „Elastizität“ dieser Verteilungsfamilie im Hinblick auf die Gestalt (Rechtsschiefe) und die Skalierung. Es kann also davon ausgegangen werden, daß durch diese Verteilungsfamilie eine große Anzahl in der Praxis vorkommender rechtsschiefer Verteilungen — Verteilungen von Zählgrößen nach Stetigkeitskorrektur oder Verteilungen von physikalischen Größen mit definierter unterer Grenze, um nur zwei Beispiele zu nennen — adäquat modelliert werden kann. Der Begriff Robustheit im Titel der Arbeit bezieht sich auf eben diese „Anpassungsfähigkeit“ und beinhaltet nicht, daß Abweichungen von dieser Modellannahme Gegenstand der folgenden Untersuchungen sind — im Gegenteil, die Eigenschaften der Familie der Gammaverteilungen werden in dem nachfolgend vorgestellten Verfahren maßgeblich ausgenutzt werden.

Überblick

Nach der Vorstellung einer Reihe später benötigter Eigenschaften des Gammaverteilungsmodells in Kapitel 2 werden in Kapitel 3 einige numerische Verfahren zur Berechnung der zur Konstruktion benötigten Verteilungen diskutiert. In Kapitel 4 wird schließlich der zur Konstruktion des Konfidenzintervalls gewählte Ansatz vorgestellt. Ziel ist dabei — wie oben bereits erwähnt — eine möglichst vollständige

Ausschöpfung der in diesem Modell enthaltenen Information, um dem Idealziel der Konstruktion eines einseitigen Konfidenzintervalls für die Varianz möglichst nahe zu kommen. Die Familie der Gammaverteilungen kann dabei jedoch nicht in der üblichen Weise durch Gestalt- und Skalierungsparameter parametrisiert werden, sondern eben durch die Varianz und den Gestaltparameter. Da in dieser Parametrisierung die Gammaverteilungen keine Exponentialgestalt mehr besitzen und somit wesentliche für die Konstruktion ähnlicher Konfidenzintervalle notwendige analytische Eigenschaften verloren gehen, müssen diese durch numerische Betrachtungen „ersetzt“ werden.

Nach der Vorstellung der Zielsetzung, der Darstellung und Motivierung der Konstruktionsvorschrift für das Intervall, die in einem iterativen Verfahren eine schließlich immer „bessere“ Annäherung des Konfidenzniveaus an endlich vielen — in ihrer Zahl im wesentlichen durch die Rechnerkapazität beschränkten — Stützstellen bewerkstelligt, werden einige Beispiele solchermaßen konstruierter Konfidenzintervalle angegeben. Dabei wird eine Darstellung des Konfidenzintervalls angegeben, die eine einfache Tabellierung ermöglicht. Da die Konstruktionsvorschrift für das Konfidenzintervall die näherungsweise Erreichung des Konfidenzniveaus nur an endlich vielen Stützstellen gewährleisten kann, wird in Kapitel 5 schließlich eine theoretische Abschätzung der Überdeckungswahrscheinlichkeiten zwischen diesen Stützstellen entwickelt und an einigen Beispielen demonstriert.

Sämtliche Berechnungen wurden an einem Personal Computer in der Programmiersprache PASCAL durchgeführt. Die Computerzeichnungen wurden mit Hilfe des Graphiksystems des Leibniz-Rechenzentrums der Bayerischen Akademie der Wissenschaften erstellt.

Bei Herrn Prof. Dr. K. Weichselberger möchte ich mich an dieser Stelle herzlich für sein großes und fortwährendes Interesse sowie für die in zahlreichen Diskussionen vermittelten überaus wertvollen Anregungen bedanken.

Mein besonderer Dank gilt auch Herrn Prof. Dr. B. Rüger für die freundliche Bereitschaft, das Zweitreferat zu übernehmen.

Bedanken möchte ich mich auch bei meinen Kolleginnen und Kollegen Frau Dr. S. Pöhlmann, Frau Dr. C. Schneider und Herrn Dr. D. Schremmer für die in langen und ergiebigen Gesprächen gewonnenen Erkenntnisse.

Frau Diplom-Statistikerin A. Rösch und Herrn Dr. H. Schmidbauer möchte ich meinen Dank für zahlreiche Hilfestellungen bei der Erstellung der Abbildungen aussprechen.

2. Grundlagen

In diesem Kapitel sollen einige Grundlagen des Gammaverteilungsmodells vorgestellt werden, auf die in späteren Kapiteln dann zurückgegriffen werden kann. Da die Zusammenstellung der Eigenschaften ausschließlich unter diesem Aspekt geschieht und nicht unter der Zielsetzung einen umfassenden Überblick über die Familie der Gammaverteilungen zu bieten, mag es so erscheinen, daß einigen weniger wichtigen Eigenschaften ein zu großes Gewicht beigemessen wird, während vermeintlich wichtigere nicht erwähnt werden. Eine sehr umfangreiche Übersicht über das Gammaverteilungsmodell ist bei Johnson und Kotz (1970, S. 166ff) zu finden. Eine ausführliche Beschreibung des Schätzproblems in der Gammaverteilungssituation bieten Bowman und Shenton (1988).

2.1. Die Gammaverteilung

In diesem Abschnitt sollen die Familie der Gammaverteilungen sowie einige wesentliche im weiteren benötigte Eigenschaften vorgestellt werden.

Definition

Sei X eine positive Zufallsvariable, deren Verteilung durch die Dichtefunktion

$$f_{a,b}(x) = \frac{1}{\Gamma(a)b^a} x^{a-1} e^{-\frac{x}{b}}, \quad a > 0, b > 0 \quad (2.1)$$

bestimmt ist. Die Verteilung von X heißt Gammaverteilung, der Parameter a Gestaltparameter (shape-parameter) und der Parameter b Skalierungsparameter (scale-parameter). Die Aussage: „ X ist gammaverteilt mit Parametern a und b “ soll durch die Schreibweise

$$X \sim G(a, b)$$

zum Ausdruck gebracht werden.

Spezialfälle

Bevor die Gestalt der Gammaverteilungsdichte näher betrachtet wird, sollen einige Spezialfälle der Gammaverteilung angegeben werden.

I **Die χ^2 -Verteilung:** Für $a = \frac{k}{2}$ mit $k \in \mathbb{N}$ und $b = 2$ ergibt sich die Dichtefunktion

$$f_{\frac{k}{2},2}(x) = \frac{1}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$

der χ^2 -Verteilung mit k Freiheitsgraden.

II **Die Exponentialverteilung:** Die Exponentialverteilung mit Dichtefunktion

$$f(x) = \lambda e^{-\lambda x}$$

ist eine Gammaverteilung mit Parametern $a = 1$ und $b = \frac{1}{\lambda}$.

III **Die Erlang-Verteilung:** Die Summe von k unabhängigen, identisch mit Parameter λ exponentialverteilten Zufallsvariablen ist Erlang- k -verteilt und besitzt die Dichtefunktion

$$f(x) = \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}.$$

Die Erlang- k -Verteilung ist somit die Gammaverteilung mit Parametern $a = k$ und $b = \frac{1}{\lambda}$.

Die Gestalt der Gammaverteilungsdichten

In Abb. 2.1 sind einige Vertreter aus der Schar der Gammaverteilungsdichten gezeichnet. Zusätzlich ist noch die Dichtefunktion der Normalverteilung mit Erwartungswert $\mu = 10b$ und Varianz $\sigma^2 = 10b^2$ eingezeichnet, deren Bedeutung an späterer Stelle erläutert werden wird.

Die folgenden Aussagen zur Gestalt der Dichtefunktion lassen sich größtenteils anhand der Abb. 2.1 nachvollziehen. Auf die Beweise, die sämtlich elementar sind, soll hier verzichtet werden.

$a < 1$: $f_{a,b}$ ist streng antiton mit $\lim_{x \rightarrow 0} f_{a,b}(x) = \infty$ (Beispiel: χ^2 -Verteilung mit einem Freiheitsgrad).

$a = 1$: $f_{1,b}$ ist streng antiton mit $\lim_{x \rightarrow 0} f_{1,b}(x) = \frac{1}{b}$ (Exponentialverteilung).

$1 < a < 2$: $f_{a,b}$ ist im Bereich $x \in (0, b(a-1))$ streng isoton mit $\lim_{x \rightarrow 0} f_{a,b}(x) = 0$, besitzt an der Stelle $x = b(a-1)$ das Maximum $f_{a,b}(b(a-1)) = \frac{1}{b\Gamma(a)} (a-1)^{a-1} e^{-(a-1)}$ und ist im Bereich $x > b(a-1)$ streng antiton. Für die Steigung der Dichtefunktion gilt der Grenzübergang $\lim_{x \rightarrow 0} \frac{\partial}{\partial x} f_{a,b}(x) = \infty$.

$a = 2$: Es gelten die gleichen Aussagen wie im Fall $1 < a < 2$, lediglich der oben angegebene Grenzwert der Steigung lautet $\lim_{x \rightarrow 0} \frac{\partial}{\partial x} f_{2,b}(x) = \frac{1}{b^2}$ (Beispiel: χ^2 -Verteilung mit 4 Freiheitsgraden).

$a > 2$: Es gilt $\lim_{x \rightarrow 0} \frac{\partial}{\partial x} f_{a,b}(x) = 0$. Ansonsten gelten die gleichen Aussagen wie für $1 < a \leq 2$.

Für $a \rightarrow 0$ gilt aufgrund des Grenzwertes $\lim_{a \rightarrow 0} \frac{1}{\Gamma(a)} = 0$ für alle $x > 0$:

$$\lim_{a \rightarrow 0} f_{a,b}(x) = 0$$

Die Momente der Gammaverteilung

Das r -te Moment existiert für alle reellen $r > -a$ und lautet

$$E(X^r) = \frac{\Gamma(a+r)}{\Gamma(a)} b^r, \quad (2.2)$$



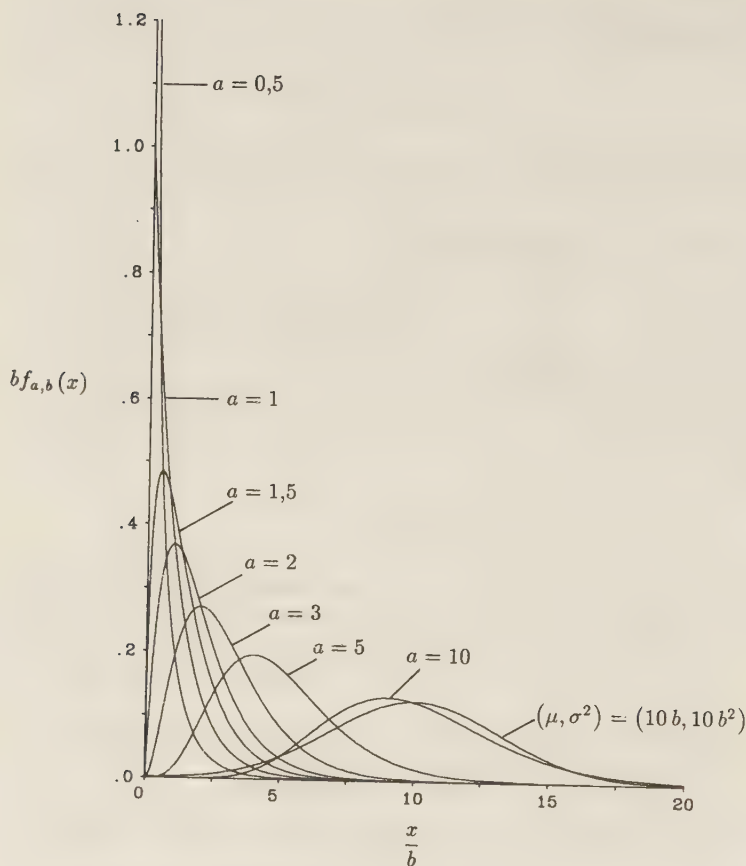


Abb. 2.1 Einige Gammaverteilungsdichten
wie aus der folgenden Umformung leicht zu erkennen ist:

$$\begin{aligned}
 E(X^r) &= \int_0^{\infty} x^r \frac{1}{\Gamma(a) b^a} x^{a-1} e^{-\frac{x}{b}} dx \\
 &= \frac{\Gamma(a+r)}{\Gamma(a)} b^r \int_0^{\infty} \frac{1}{\Gamma(a+r) b^{a+r}} x^{a+r-1} e^{-\frac{x}{b}} dx.
 \end{aligned}$$

Der Integrand des zweiten Integrales ist die Dichtefunktion der Gammaverteilung mit den Parametern $a + r$ und b , das Integral besitzt also den Wert 1.

Damit lassen sich leicht die folgenden Momente berechnen:

$$\begin{aligned}\mu &\stackrel{\text{def}}{=} E(X) = ab \\ \sigma^2 &\stackrel{\text{def}}{=} \text{Var}(X) = ab^2 \\ \mu_3 &\stackrel{\text{def}}{=} E((X - \mu)^3) = 2ab^3 \\ \mu_4 &\stackrel{\text{def}}{=} E((X - \mu)^4) = 3(a + 2)ab^4,\end{aligned}\tag{2.3}$$

woraus sich für den Variationskoeffizienten, die Schiefe und die Kurtosis

$$\begin{aligned}\frac{\sigma}{\mu} &= \frac{1}{\sqrt{a}} \\ \frac{\mu_3}{\sigma^3} &= \frac{2}{\sqrt{a}} \\ \frac{\mu_4}{\sigma^4} &= 3 + \frac{6}{a}\end{aligned}$$

ergeben. Insbesondere ist zu erkennen, daß $\frac{\mu_3}{\sigma^3}$ und $\frac{\mu_4}{\sigma^4}$ für $a \rightarrow \infty$ gegen die entsprechenden Werte der Normalverteilung konvergieren. Dieses bestätigt das bekannte und über die Konvergenz der charakteristischen Funktionen beweisbare Resultat, daß für festes $b > 0$ die Verteilung von $\frac{X - \mu}{\sigma}$ für $a \rightarrow \infty$ gegen die Standardnormalverteilung konvergiert. Um diese Näherung zu verdeutlichen, ist in Abb. 2.1 zusätzlich zu der Gammaverteilungsdichte mit $a = 10$ noch die Dichtefunktion der Normalverteilung mit Erwartungswert $10b$ und Varianz $10b^2$ eingezeichnet. Anstelle der sonst nicht mehr benötigten charakteristischen Funktion der Gammaverteilung soll hier die Laplace-Transformierte angegeben werden.

Laplace-Transformierte

Die Laplace-Transformierte der Dichte der Gammaverteilung lautet für $r > -\frac{1}{b}$:

$$E(e^{-rX}) = (1 + rb)^{-a},\tag{2.4}$$

wie sich analog zur Bestimmung der Momente zeigen läßt:

$$\begin{aligned}E(e^{-rX}) &= \int_0^{\infty} e^{-rx} f_{a,b}(x) dx \\ &= (1 + rb)^{-a} \int_0^{\infty} \frac{1}{\Gamma(a) \left(\frac{b}{1+rb}\right)^a} x^{a-1} e^{-\left(\frac{x}{1+rb}\right)} dx.\end{aligned}$$

Unter dem Integral steht diesmal die Dichtefunktion der Gammaverteilung mit den Parametern a und $\frac{b}{1+rb}$.

Additionstheorem

Mit Hilfe der Laplace-Transformierten lassen sich unmittelbar die folgenden Aussagen beweisen:

I Sind X und Y unabhängige, gammaverteilte Zufallsvariable mit

$$X \sim G(a_x, b) \quad \text{und} \quad Y \sim G(a_y, b)$$

so gilt

$$X + Y \sim G(a_x + a_y, b).$$

II Ist X gammaverteilt

$$X \sim G(a, b),$$

dann gilt

$$\forall d > 0: \quad dX \sim G(a, db).$$

Die Eigenschaft II erklärt die Funktion des Parameters b , die auch der Achsenbeschriftung der Abb. 2.1 zugrundeliegt. Für $d = \frac{1}{b}$ ergibt sich $\frac{X}{b} \sim G(a, 1)$, der Parameter b ist also ein Skalierungsparameter. Das obige Theorem erlaubt auch die Bestimmung der Verteilung des arithmetischen Mittels von gammaverteilten Zufallsvariablen:

Sind $(X_1, \dots, X_n)$ unabhängig, identisch $G(a, b)$ -verteilte Zufallsvariable, so gilt:

$$\bar{X} \sim G\left(an, \frac{b}{n}\right).$$

Zur Notation

Da im weiteren ein Konfidenzintervall für die Varianz σ^2 der Gammaverteilung bestimmt werden soll, wird häufig die — in der Literatur sonst unübliche — Parametrisierung der Gammaverteilung durch die Parameter a und σ^2 vorgenommen. Da (a, σ^2) aus (a, b) durch eine bijektive Transformation hervorgeht, sind beide Parametrisierungen äquivalent. Demzufolge wird häufig auch die Schreibweise

$$X \sim G(a, \sigma^2) \quad \text{und} \quad \bar{X} \sim G\left(an, \frac{\sigma^2}{n}\right)$$

verwendet. Damit ist strenggenommen eine uneinheitliche Verwendung des Symbols G verbunden, doch soll diese Doppeldeutigkeit in Kauf genommen werden, um keine zu komplizierte Notation zu erhalten, zumal aus dem Zusammenhang heraus immer klar ersichtlich ist, welche Bedeutung der zweite Parameter (Skalierungsparameter oder Varianz) besitzt. Die Verteilungsfunktion der Gammaverteilung wird dagegen einheitlich mit dem Symbol

$$\mathcal{G}_{a, \sigma^2}$$

bezeichnet werden, was allerdings zur Folge hat, daß die an manchen Stellen benötigte Verteilungsfunktion der χ^2 -Verteilung mit k Freiheitsgraden in der ungewöhnlichen Darstellung

$$\mathcal{G}_{\frac{k}{2}, 2k}$$

angegeben wird.

Die Gestalt der $G(a, \sigma^2)$ -Verteilung

In Abb. 2.2 sind einige Verteilungen aus der Schar der Gammaverteilungen mit Gestaltparameter a und Skalierungsparameter $b = \frac{1}{\sqrt{a}}$ — also mit $\sigma^2 = 1$ — eingezeichnet.

2.2. Schätzer im Gammaverteilungsmodell

Exponentialfamilie

Aus der Darstellung

$$f_{a,b}(x) = \frac{1}{\Gamma(a)b^a} e^{(a-1)\ln x - \frac{x}{b}}$$

der Dichtefunktion der Gammaverteilung ist zu erkennen, daß die Familie der Gammaverteilungen eine Exponentialfamilie ist.

Die Dichtefunktion einer Stichprobe $(X_1, \dots, X_n)$ von unabhängigen, identisch $G(a, b)$ -verteilten Zufallsvariablen lautet

$$\begin{aligned} f_{a,b}(x_1, \dots, x_n) &= \frac{1}{\Gamma(a)^n b^{an}} \left(\sqrt[n]{\prod_{i=1}^n x_i} \right)^{n(a-1)} e^{-n \frac{x}{b}} \\ &= \frac{1}{\Gamma(a)^n b^{an}} e^{n(a-1) \ln \left(\sqrt[n]{\prod_{i=1}^n x_i} \right) - \frac{n}{b} \bar{x}}. \end{aligned} \quad (2.5)$$

Aus dieser Darstellung folgt, daß die Statistik $\left(\bar{X}, \ln \left(\sqrt[n]{\prod_{i=1}^n X_i} \right) \right)$ suffizient für die Klasse der Gammaverteilungen ist.

Maximum-Likelihood-Schätzer für a und σ^2

Parametrisiert durch (a, σ^2) lautet die obige Dichtefunktion:

$$f_{a,\sigma^2}(x_1, \dots, x_n) = \frac{(\sqrt{a})^{an}}{\Gamma(a)^n (\sqrt{\sigma^2})^{an}} e^{n(a-1) \ln \left(\sqrt[n]{\prod_{i=1}^n x_i} \right) - \frac{n\sqrt{a}}{\sqrt{\sigma^2}} \bar{x}}.$$

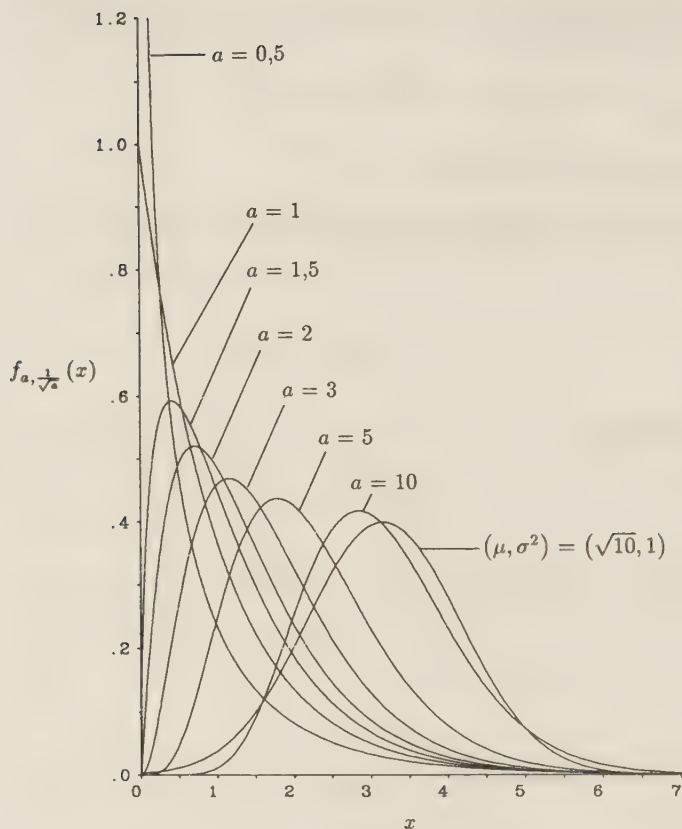


Abb. 2.2 Gammaverteilungsdichten mit $\sigma^2 = 1$

Durch Ableiten von $\ln f_{a, \sigma^2}$ nach a und σ^2 ergibt sich

$$\begin{aligned} \frac{\partial}{\partial a} \ln f_{a, \sigma^2}(x_1, \dots, x_n) &= \\ n \left[-\frac{\Gamma'(a)}{\Gamma(a)} - \frac{\ln(\sigma^2)}{2} + \frac{\ln(a)}{2} + \frac{1}{2} + \ln \left(\sqrt[n]{\prod_{i=1}^n x_i} - \frac{\bar{x}}{2\sqrt{a\sigma^2}} \right) \right] \\ \frac{\partial}{\partial \sigma^2} \ln f_{a, \sigma^2}(x_1, \dots, x_n) &= \frac{\sqrt{a}}{2(\sqrt{\sigma^2})^3} [-\sqrt{a\sigma^2} - \bar{x}]. \end{aligned}$$

Die zweite Gleichung liefert durch Nullsetzen

$$\bar{X} = \sqrt{\hat{a}_{ML} \hat{\sigma}_{ML}^2}$$

und durch Nullsetzen der ersten Gleichung

$$\ln \left(\sqrt[n]{\prod_{i=1}^n X_i} \right) = \frac{\Gamma'(\hat{a}_{ML})}{\Gamma(\hat{a}_{ML})} + \frac{\ln(\hat{\sigma}_{ML}^2)}{2} - \frac{\ln(\hat{a}_{ML})}{2}.$$

Durch Subtraktion von $\ln \bar{X}$ ergibt sich schließlich

$$\ln \left(\sqrt[n]{\prod_{i=1}^n X_i} \right) - \ln \bar{X} = \frac{\Gamma'(\hat{a}_{ML})}{\Gamma(\hat{a}_{ML})} - \ln(\hat{a}_{ML}).$$

Die Funktion $\psi(x) \stackrel{\text{def}}{=} \frac{\partial}{\partial x} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ heißt Psi-Funktion oder auch Digamma-

Funktion, der Quotient aus geometrischem und arithmetischem Mittel $\frac{\sqrt[n]{\prod_{i=1}^n X_i}}{\bar{X}}$, soll im folgenden stets mit

$$Q \stackrel{\text{def}}{=} \frac{\sqrt[n]{\prod_{i=1}^n X_i}}{\bar{X}}, \quad (2.6)$$

bezeichnet werden, womit sich schließlich die beiden Bestimmungsgleichungen

$$\begin{aligned} \ln Q &= \psi(\hat{a}_{ML}) - \ln(\hat{a}_{ML}) \\ \hat{\sigma}_{ML}^2 &= \frac{\bar{X}^2}{\hat{a}_{ML}} \end{aligned} \quad (2.7)$$

ergeben. Die Zufallsvariable Q nimmt nur Werte im Intervall $(0,1)$ an, ihre Verteilung wird im nächsten Abschnitt betrachtet werden. Bevor der funktionale Zusammenhang zwischen $\hat{a}_{ML}$ und Q näher betrachtet wird, bleibt nur noch festzustellen, daß mit $\left(\bar{X}, \ln \left(\sqrt[n]{\prod_{i=1}^n X_i} \right) \right)$ auch die Statistik $(\hat{a}_{ML}, \hat{\sigma}_{ML}^2)$ suffizient ist (die Abbildung von Q auf $\hat{a}_{ML}$ ist bijektiv, wie im folgenden gezeigt wird). Bei der Herleitung des Konfidenzintervalls für σ^2 in Kapitel 4 wird dann stets das Tupel $(\hat{a}_{ML}, \hat{\sigma}_{ML}^2)$ verwendet werden. Die Angabe dieser Statistik hat gegenüber der ursprünglichen insbesondere den Vorteil, daß für den mittels Bereichsschätzung zu schätzenden Parameter σ^2 zusätzlich ein Punktschätzer angegeben wird, außerdem ist der durch $\hat{a}_{ML}$ geschätzte Parameter a gut interpretierbar.

2.3. Die Verteilung von Q

In diesem Abschnitt sollen die Dichtefunktion sowie die Momente der Zufallsvariablen Q betrachtet werden. Da in einer Stichprobe vom Umfang $n = 1$ das geometrische und das arithmetische Mittel stets identisch sind, besitzt die mit Wahrscheinlichkeit 1 definierte Zufallsvariable Q in diesem Fall stets den Wert 1. Daher wird im folgenden — ohne dieses jedesmal explizit anzugeben — stets $n \geq 2$ vorausgesetzt werden.

Skalierungsinvarianz von Q

Aus der Beziehung

$$\forall d > 0: \quad \frac{\sqrt[n]{\prod_{i=1}^n x_i}}{\frac{1}{n} \sum_{i=1}^n x_i} = \frac{\sqrt[n]{\prod_{i=1}^n dx_i}}{\frac{1}{n} \sum_{i=1}^n dx_i}$$

ergibt sich sofort die folgende Aussage:

Der Wert der Zufallsvariablen Q ändert sich nicht, wenn alle Elemente der Stichprobe mit einer positiven Konstante d multipliziert werden:

$$\forall d > 0: \quad Q(x_1, \dots, x_n) = Q(dx_1, \dots, dx_n)$$

Die bedingte Dichte von Q gegeben $\bar{X}$

Aus der Exponentialgestalt (2.5) der Dichtefunktion einer Stichprobe $(X_1, \dots, X_n)$ folgt (s. Barndorff-Nielsen (1978, S. 126)), daß die Dichtefunktion der suffizienten Statistik

$$(U, T) \stackrel{\text{def}}{=} \left(\ln \left(\sqrt[n]{\prod_{i=1}^n X_i} \right), \bar{X} \right) \quad (2.8)$$

für $u < 0$ und $t > 0$ die Gestalt

$$f_{a,b}^{U,T}(u, t) = \frac{1}{\Gamma(a)^n b^{an}} e^{n(a-1)u - \frac{t}{b} u} p(u, t) \quad (2.9)$$

besitzt, wobei aus der Theorie der Exponentialfamilien über die Funktion p lediglich bekannt ist, daß sie nicht von den Parametern a und b abhängt (näheres zu einer Transformation der sehr komplizierten Funktion p an späterer Stelle). Die gemeinsame Dichte von Q und T kann für $0 < q < 1$ und $t > 0$ mittels einer einfachen Dichtetransformation berechnet werden:

$$f_{a,b}^{Q,T}(q, t) = \frac{1}{\Gamma(a)^n b^{an}} q^{n(a-1)-1} t^{n(a-1)} e^{-\frac{t}{b} \ln(qt)} p(\ln(qt), t) \quad (2.10)$$

Da $T = \bar{X}$ gammaverteilt ist mit den Parametern an und $\frac{b}{n}$, lautet die bedingte Dichte von Q unter $T = t$ für $0 < q < 1$ und $t > 0$

$$\begin{aligned} f_a^{Q|T=t}(q) &= \frac{f_{a,b}^{Q,T}(q, t)}{f_{an, \frac{b}{n}}^T(t)} \\ &= \frac{\Gamma(an)}{n^{an-1} \Gamma(a)^n} q^{an-1} \frac{1}{nq} (qt)^{1-n} p(\ln(qt), t). \end{aligned}$$

Nach Einführung der Funktion

$$\tilde{p}(q, t) \stackrel{\text{def}}{=} \frac{1}{nq} (qt)^{1-n} p(\ln(qt), t)$$

ergibt sich schließlich die Darstellung

$$f_a^{Q|T=t}(q) = \frac{\Gamma(an)}{n^{an-1} \Gamma(a)^n} q^{an-1} \tilde{p}(q, t). \quad (2.11)$$

Diese Darstellung zeigt, daß die bedingte Verteilung nur noch vom Parameter a abhängt.

Momente und Dichtefunktion von Q

Die bedingten Momente von Q gegeben $T = t$ lassen sich für $r > -an$ berechnen:

$$\begin{aligned}
 E(Q^r | T = t) &= \int_0^1 q^r \frac{\Gamma(an)}{n^{an-1} \Gamma(a)^n} q^{an-1} \tilde{p}(q, t) dq \\
 &= \frac{\Gamma(an) \Gamma(a + \frac{r}{n})^n}{\Gamma(a)^n \Gamma(an+r)} n^r \int_0^1 \frac{\Gamma(an+r)}{n^{an+r-1} \Gamma(a + \frac{r}{n})^n} q^{an+r-1} \tilde{p}(q, t) dq \\
 &= \frac{\Gamma(an) \Gamma(a + \frac{r}{n})^n}{\Gamma(a)^n \Gamma(an+r)} n^r \quad (2.12)
 \end{aligned}$$

Die Momente hängen nicht von t ab, woraus — da die Zufallsvariable Q einen beschränkten Träger besitzt und ihre Verteilung deshalb durch ihre Momente eindeutig bestimmt wird (s. Feller (1971, S. 224ff)) — folgt, daß die Zufallsvariablen Q und $\bar{X}$ unabhängig sind. Zudem folgt, daß die Funktion $\tilde{p}$ nur von q , nicht aber von t abhängt, womit die Dichtefunktion von Q für $0 < q < 1$

$$f_a^Q(q) = \frac{\Gamma(an)}{n^{an-1} \Gamma(a)^n} q^{an-1} \tilde{p}(q) \quad (2.13)$$

lautet. Die Momente der Verteilung von Q sowie die Unabhängigkeit der Zufallsvariablen Q und $\bar{X}$ sind ein wohlbekanntes Ergebnis. Die Unabhängigkeit der Variablen wurde bereits von Nair (1938, S. 287ff) bewiesen. Cox und Lewis (1966) geben eine entsprechende Aussage für die Verteilungen von $\hat{a}_{ML}$ und $\bar{X}$ an. Ein im Rahmen der allgemeinen Theorie der Exponentialfamilien geführter Beweis über das Theorem von Basu (Basu (1955)), findet sich bei Barndorff-Nielsen (1978, S. 183).

Eine andere Darstellung der Verteilung von Q

Im folgenden soll ein Zusammenhang zwischen der Verteilung von Q und der Verteilung des Produktes von unabhängigen, nicht identisch betaverteilten Zufallsvariablen angegeben werden. Dieser Zusammenhang wird durch einen Vergleich der Momente ersichtlich. Zu diesem Zwecke werden in einem ersten Schritt die Momente von Q in einer anderen Form dargestellt.

Die Integrationskonstante $\frac{\Gamma(an)}{n^{an-1} \Gamma(a)^n}$ kann unter zweifacher Anwendung der für $m \in \mathbb{N}$ gültigen Gaußschen Multiplikationsformel (s. Abramowitz, Stegun (1984, S. 77))

$$\Gamma(mz) = m^{mz - \frac{1}{2}} (2\pi)^{-\frac{m-1}{2}} \prod_{k=0}^{m-1} \Gamma\left(z + \frac{k}{m}\right) \quad (2.14)$$

umgeformt werden:

Für $m = n$ und $z = a$ ergibt sich

$$\frac{\Gamma(an)}{n^{an-1}\Gamma(a)^n} = n^{\frac{1}{2}} (2\pi)^{-\frac{n-1}{2}} \prod_{k=1}^{n-1} \frac{\Gamma(a + \frac{k}{n})}{\Gamma(a)} \quad (2.15)$$

während sich aus $m = n$ und $z = 1$ die Beziehung

$$1 = (2\pi)^{\frac{n-1}{2}} n^{-\frac{1}{2}} \frac{1}{\prod_{k=1}^{n-1} \Gamma(\frac{k}{n})} \quad (2.16)$$

bestätigen läßt. Durch Multiplikation der rechten Seite der Gleichung (2.15) mit der durch die rechte Seite von (2.16) dargestellten Konstanten 1 ergibt sich schließlich unter Einführung der Betafunktion

$$\frac{\Gamma(an)}{n^{an-1}\Gamma(a)^n} = \prod_{k=1}^{n-1} \frac{1}{B(a, \frac{k}{n})}. \quad (2.17)$$

Da das r -te Moment $E(Q^r)$ sich nach (2.12) als Quotient zweier Integrationskonstanten darstellen läßt, ergibt sich

$$\forall r > -an : \quad E(Q^r) = \prod_{k=1}^{n-1} \frac{B(a + \frac{r}{n}, \frac{k}{n})}{B(a, \frac{k}{n})}. \quad (2.18)$$

Die Momente einer mit Parametern α und β ($\alpha > 0, \beta > 0$) betaverteilten Zufallsvariablen Y , deren Dichtefunktion für $0 < y < 1$ durch

$$f_{\alpha, \beta}(y) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}$$

gegeben ist, lauten für $r > -\alpha$

$$E(Y^r) = \frac{B(\alpha + r, \beta)}{B(\alpha, \beta)}. \quad (2.19)$$

Aus der bereits erwähnten Eigenschaft, daß eine auf einem beschränkten Träger definierte Verteilung eindeutig durch ihre Momente charakterisiert wird, folgt aus der Identität der Momente (2.18) von Q und der aus (2.19) berechenbaren Momente der im folgenden definierten Variablen $\sqrt[n]{\prod_{k=1}^{n-1} Y_k}$:

Seien $Y_1, \dots, Y_{n-1}$ unabhängige, betaverteilte Zufallsvariable, mit $Y_k \sim \text{Beta}(a, \frac{k}{n})$, $k = 1, \dots, n-1$, so gilt:

$$\sqrt[n]{\prod_{k=1}^{n-1} Y_k} \quad \text{und} \quad Q \quad \text{sind identisch verteilt.} \quad (2.20)$$

Die Funktion $\tilde{p}$

Aufgrund der Beziehung (2.20) läßt sich nun auch die Funktion $\tilde{p}$ berechnen:

$n = 2$: Q^2 ist betaverteilt mit Parametern a und $\frac{1}{2}$. Der Transformationssatz für Dichten liefert damit die Dichtefunktion von Q :

$$f_a^Q(q) = 2 \frac{\Gamma(a + \frac{1}{2})}{\Gamma(a) \Gamma(\frac{1}{2})} q^{2(a-1)+1} (1-q^2)^{-\frac{1}{2}}.$$

Unter Anwendung von (2.14) und $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ ergibt sich schließlich aus (2.13)

$$\tilde{p}(q) = \frac{2}{\sqrt{1-q^2}}.$$

Diese Darstellung der Dichtefunktion für den Fall $n = 2$ wurde bereits von Nayer (1936, S. 43) angegeben.

$n > 2$: Die Dichtefunktion des oben definierten Zufallsvektors $Y_1, \dots, Y_{n-1}$ lautet

$$f_a^Y(y_1, \dots, y_{n-1}) = \frac{1}{\prod_{k=1}^{n-1} B(a, \frac{k}{n})} \prod_{k=1}^{n-1} y_k^{a-1} (1-y_k)^{\frac{k}{n}-1},$$

$0 < y_k < 1, k = 1, \dots, n-1.$

Durch die Transformation $Q = \sqrt{\prod_{k=1}^{n-1} Y_k}$ ergibt sich mittels der Dichtetransformation die gemeinsame Dichte des Zufallsvektors $(Y_1, \dots, Y_{n-2}, Q)$, die im Bereich $0 < y_k < 1, k = 1, \dots, n-2, 0 < q < \sqrt[n]{\prod_{k=1}^{n-2} y_k}$ definiert ist und durch „herausintegrieren“ von $Y_1, \dots, Y_{n-2}$ die Dichtefunktion von Q :

$$f_a^Q(q) = \frac{1}{\prod_{k=1}^{n-1} B(a, \frac{k}{n})} q^{an-1} \cdot n \int \dots \int_D \prod_{k=1}^{n-2} \frac{(1-y_k)^{\frac{k}{n}-1}}{y_k} \left(1 - \frac{q^n}{\prod_{k=1}^{n-2} y_k}\right)^{-\frac{1}{n}} dy_1 \dots dy_{n-2}$$

mit

$$D = \left\{ (y_1, \dots, y_{n-2}) : 0 < y_k < 1, k = 1, \dots, n-2, q < \sqrt[n]{\prod_{k=1}^{n-2} y_k} \right\}.$$

Aus (2.13) und (2.17) ergibt sich damit für $0 < q < 1$

$$\tilde{p}(q) = n \int \dots \int_D \prod_{k=1}^{n-2} \frac{(1-y_k)^{\frac{k}{n}-1}}{y_k} \left(1 - \frac{q^n}{\prod_{k=1}^{n-2} y_k}\right)^{-\frac{1}{n}} dy_1 \dots dy_{n-2},$$

mit dem oben angegebenen D .

Die Funktion $\tilde{p}$ soll nicht weiter untersucht werden. Aus einer anderen Vorgehensweise heraus wurde bereits von Nair (1938) eine Reihenentwicklung dieser Funktion angegeben, die zeigt, daß Berechnungen der Verteilungsfunktion von Q , die von Bishop und Nair (1939) aufgrund dieser Methode vorgenommen wurden, sehr schwierig sind und mit wachsendem Stichprobenumfang n zunehmend aufwendiger werden, da $\tilde{p}$ eine äußerst komplexe Gestalt besitzt. Die bei Nair (1938, S. 283) angegebene Entwicklung lautet:

$$\tilde{p}(q) = \frac{1}{(n-2)!} \sum_{i=0}^{\infty} \frac{(in)!}{(i!)^n n^{ni}} \Delta(i, n) q^{in},$$

wobei $\Delta(i, n)$ die Determinante einer $(n-2) \times (n-2)$ Matrix ist, in die die ersten $(n-2)$ Ableitungen der ψ -Funktion, sowie die ψ -Funktion, an jeweils mehreren Stellen ausgewertet, eingehen.

Eine Darstellung der Dichtefunktion f_a^G des geometrischen Mittels findet sich bei Kullback (1934):

$$f_a^G(g) = \frac{ng^{an-1}}{\Gamma(n)\Gamma(a)^n} \sum_{i=0}^{\infty} (-1)^{n+ni+1} \left[\frac{\partial^{n-1}}{\partial t^{n-1}} \frac{g^{nt}}{\Gamma(t+1)} \right]_{t=i}.$$

Auch diese Darstellung ist ähnlich komplex wie die von Nair angegebene Darstellung der Dichtefunktion von Q .

Um die später benötigten Verteilungsfunktionen von Q und $\hat{a}_{ML}$ berechnen zu können, reicht es aus zu wissen, daß die Funktion $\hat{p}$ nicht vom Parameter a abhängt. Die numerische Bestimmung der Verteilungsfunktionen wird in Kapitel 3 aufgrund der Kenntnis der Momente $E(Q^r)$, die die Verteilung von Q sowie die Verteilung der bijektiven Transformation $\hat{a}_{ML}$ eindeutig festlegen, durchgeführt.

Die asymptotischen Momente von Q

Der unter (2.20) beschriebene Zusammenhang zwischen den Verteilungen von Q und dem Produkt von betaverteilten Zufallsvariablen, wird in Kapitel 5.1 Verwendung finden. Im folgenden soll gezeigt werden, daß gilt

Ist Y eine Beta $(a(n-1), \frac{n-1}{2})$ -verteilte Zufallsvariable, so gilt:

$$Y^{a(n-1)} \text{ und } Q^{an} \text{ konvergieren für } a \rightarrow \infty \text{ nach Verteilung gegeneinander.} \quad (2.21)$$

Für $n=2$ ist die obige Aussage trivialerweise erfüllt — die Verteilungen sind für alle $a > 0$ identisch. Für $n > 2$ ergibt sich durch Einsetzen der Stirlingschen Formel

$$\ln \Gamma(x) = \left(x - \frac{1}{2}\right) \ln x - x + \frac{1}{2} \ln(2\pi) + o(1), \quad \text{für } x \rightarrow \infty$$

in

$$\ln E(Q^{anr}) = \ln \left(\frac{\Gamma(an) \Gamma(a + \frac{anr}{n})^n}{\Gamma(a)^n \Gamma(an + anr)} n^{anr} \right)$$

der für alle $r > -1$ definierte Grenzwert

$$\lim_{a \rightarrow \infty} E(Q^{anr}) = (1+r)^{-\frac{n-1}{2}}. \quad (2.22)$$

Analog ergibt sich für das r -te Moment von $Y^{a(n-1)}$

$$E(Y^{a(n-1)r}) = \frac{B(a(n-1)(1+r), \frac{n-1}{2})}{B(a(n-1), \frac{n-1}{2})}$$

der Grenzwert

$$\lim_{a \rightarrow \infty} E(Y^{a(n-1)r}) = (1+r)^{-\frac{n-1}{2}}.$$

Dieses bedeutet aber, da, wie bereits erwähnt, die Momente einer auf einem beschränkten Träger definierten Verteilung diese eindeutig bestimmen, die Verteilungskonvergenz.

Im Falle des Grenzüberganges $a \rightarrow 0$ existiert keine vergleichbare Aussage. Der für $r > -1$ definierte Grenzwert von $E(Q^{anr})$ läßt sich unter Verwendung der für $x > 0$ gültigen Rekursionsbeziehung

$$\Gamma(x+1) = x\Gamma(x)$$

und des Grenzwertes $\lim_{x \rightarrow 0} \Gamma(x+1) = 1$ in einfacher Weise berechnen:

$$\lim_{a \rightarrow 0} E(Q^{anr}) = (1+r)^{-(n-1)}. \quad (2.23)$$

Für die Momente von $Y^{a(n-1)}$ ergibt sich demgegenüber

$$\lim_{a \rightarrow 0} E(Y^{a(n-1)r}) = (1+r)^{-1}.$$

Sie stimmen damit lediglich für $n = 2$ mit den asymptotischen Momenten von Q^{an} überein.

2.4. Die Verteilung von $\hat{a}_{ML}$

Zwischen den Zufallsvariablen Q und $\hat{a}_{ML}$ wurde in Abschnitt 2.2 der Zusammenhang

$$\ln Q = \psi(\hat{a}_{ML}) - \ln(\hat{a}_{ML}) \quad (2.24)$$

hergeleitet.

Die Funktion ρ

Es soll gezeigt werden, daß die für $x > 0$ definierte Funktion ρ :

$$\rho(x) \stackrel{\text{def}}{=} \psi(x) - \ln(x) \quad (2.25)$$

eine streng isotone, stetige Funktion ist mit den Grenzwerten

$$\lim_{x \rightarrow 0} \rho(x) = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \rho(x) = 0. \quad (2.26)$$

Aufgrund dieser Eigenschaft ist der Maximum-Likelihood Schätzer $\hat{a}_{ML}$ eine eindeutige Transformation der Zufallsvariablen Q und zwischen den Verteilungsfunktionen von Q und $\hat{a}_{ML}$ besteht der einfache Zusammenhang

$$\begin{aligned} P_a \{Q \leq q\} &= P_a \{\hat{a}_{ML} \leq \rho^{-1}(\ln q)\} \\ P_a \{\hat{a}_{ML} \leq \hat{a}\} &= P_a \{Q \leq e^{\rho(\hat{a})}\} \end{aligned} \quad (2.27)$$

(Die Schreibweise P_a soll andeuten, daß die Wahrscheinlichkeit vom Wert des Parameters b nicht abhängt).

Im folgenden wird eine Reihe von Eigenschaften der Digammafunktion ψ verwendet. Diese Funktion sowie ihre Ableitungen — die Polygammafunktionen — werden ausführlich von Spanier, Oldham (1987, S. 423ff) sowie Abramowitz, Stegun (1984, S. 79ff) behandelt. Die im folgenden verwendeten Formeln sind allesamt dort zu finden, ohne daß dieses jedesmal erwähnt werden wird. Die Monotonie der Funktion ρ kann durch die Bildung ihrer Ableitung gezeigt werden. Die Funktion ψ ist ebenso wie die Logarithmusfunktion im Bereich positiver Argumente eine streng isotone Funktion, mit den Grenzwerten

$$\lim_{x \rightarrow 0} \psi(x) = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \psi(x) = \infty.$$

Ihre Ableitung — die Trigammafunktion — besitzt ebenfalls die gleichen Grenzwerte wie die Ableitung der Logarithmusfunktion, so daß auf diese Weise noch keine Aussagen über die Funktion ρ möglich sind. Die Ableitung der Reihenentwicklung

$$\rho(x) = - \sum_{j=0}^{\infty} \left[\frac{1}{j+x} - \ln \left(1 + \frac{1}{j+x} \right) \right]$$

lautet

$$\rho'(x) = \sum_{j=0}^{\infty} \frac{1}{(j+x)^2} \frac{1}{j+x+1},$$

woraus die Monotonie von ρ sofort zu erkennen ist (sämtliche Summanden sind positiv).

Aus der für $x \neq 0$ gültigen Rekursionsbeziehung

$$\psi(x+1) = \psi(x) + \frac{1}{x} \quad (2.28)$$

ergibt sich für ρ

$$x\rho(x) = x\rho(x+1) - x \ln\left(\frac{x}{x+1}\right) - 1$$

und da $\rho(1) = \psi(1)$ endlich ist sowie $\lim_{x \rightarrow 0} x \ln\left(\frac{x}{x+1}\right) = 0$ gilt, folgt die Grenzwertbeziehung

$$\rho(x) = -\frac{1}{x} + o\left(\frac{1}{x}\right) \quad \text{für } x \rightarrow 0. \quad (2.29)$$

Die asymptotische Entwicklung

$$\psi(x) \sim \ln x - \frac{1}{2x} + \sum_{j=1}^{\infty} \frac{B_{2j}}{2j x^{2j}}, \quad \text{für } x \rightarrow \infty,$$

die ein Analogon zur Stirlingschen Reihe für die Ableitung des Logarithmus der Gammafunktion darstellt (B_{2j} : Bernoulli-Zahlen), führt direkt zur zweiten Grenzwertaussage

$$\rho(x) = -\frac{1}{2x} + o\left(\frac{1}{x}\right) \quad \text{für } x \rightarrow \infty. \quad (2.30)$$

In Abbildung 2.3 ist die Funktion $Q \mapsto \hat{a}_{ML} = \rho^{-1} \circ \ln Q$ gezeichnet. Aus der Darstellung

$$\rho(x) = -\frac{1}{2x} - 2 \int_0^{\infty} \frac{t}{(t^2 + x^2)(\exp(2\pi t) - 1)} dt$$

folgt — da der Integrand positiv ist — daß ρ stets kleiner als $-\frac{1}{2x}$ ist. Aus der für alle positiven x gültigen Formel

$$\psi(nx) = \ln n + \frac{1}{n} \sum_{j=0}^{n-1} \psi\left(x + \frac{j}{n}\right), \quad n = 1, 2, 3, \dots$$

folgt aufgrund der Isotonie von ψ und der Rekursionsbeziehung (2.28)

$$\psi(nx) \leq \ln n + \psi(x) + \frac{1}{x}.$$

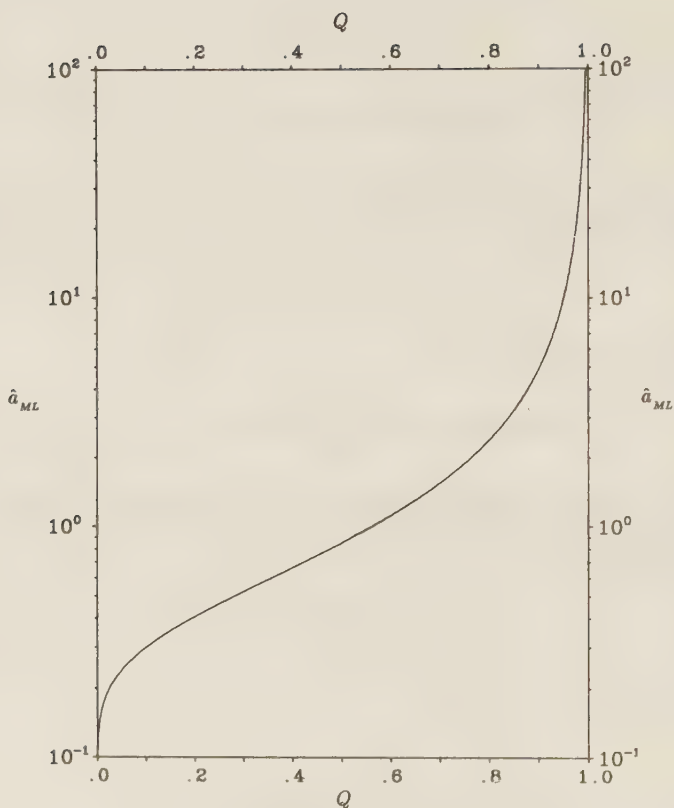


Abb. 2.3 Die Funktion: $Q \mapsto \hat{a}_{ML} = \rho^{-1} \circ \ln Q$

Damit ergibt sich die für alle positiven x erfüllte Ungleichung

$$\rho(nx) - \frac{1}{x} \leq \rho(x), \quad n = 1, 2, 3, \dots,$$

und durch den Grenzübergang $n \rightarrow \infty$ schließlich zusammen mit der ersten Abschätzung

$$\forall x > 0: \quad -\frac{1}{x} \leq \rho(x) \leq -\frac{1}{2x}. \quad (2.31)$$

Durch Einsetzen in (2.24) folgt damit die Ungleichung

$$-\frac{1}{\hat{a}_{ML}} \leq \ln Q \leq -\frac{1}{2\hat{a}_{ML}} \quad (2.32)$$

oder umgekehrt

$$-\frac{1}{2\ln Q} \leq \hat{a}_{ML} \leq -\frac{1}{\ln Q}. \quad (2.33)$$

Aus der Inversen ρ^{-1} kann die Funktion τ

$$\tau(x) \stackrel{\text{def}}{=} x\rho^{-1}(x) \quad (2.34)$$

gebildet werden, die aufgrund der Beziehung $\hat{a}_{ML} = \rho^{-1}(\ln Q)$, die Darstellung

$$\hat{a}_{ML} = \frac{\tau(\ln Q)}{\ln Q} \quad (2.35)$$

liefert. τ ist eine stetige, streng isotone und beschränkte Funktion mit

$$\lim_{q \rightarrow 0} \tau(\ln q) = -1 \quad \text{und} \quad \lim_{q \rightarrow 1} \tau(\ln q) = -\frac{1}{2}. \quad (2.36)$$

Die Dichtefunktion von $\hat{a}_{ML}$

Da die Funktion ρ die Zufallsvariablen Q und $\hat{a}_{ML}$ bijektiv aufeinander abbildet, ergibt sich die Dichtefunktion von $\hat{a}_{ML}$ für positives Argument $\hat{a}$ durch eine einfache Transformation der Dichtefunktion von Q :

$$f_{\hat{a}_{ML}}^{\hat{a}}(\hat{a}) = \frac{\Gamma(an)}{n^{an-1}\Gamma(a)^n} e^{an\rho(\hat{a})} \rho'(\hat{a}) \tilde{p}\left(e^{\rho(\hat{a})}\right). \quad (2.37)$$

2.5. Die Verteilung von $\ln Q$

In diesem Abschnitt sollen die Momente der Zufallsvariablen $\ln Q$ sowie die asymptotische Verteilung von $-2an \ln Q$ für $a \rightarrow 0$ und $a \rightarrow \infty$ angegeben werden. Vorbereitend dazu muß aber noch gezeigt werden, daß die Momente der Gammaverteilung diese eindeutig definieren.

Zur Bestimmung einer Verteilung aus ihren Momenten

In Abschnitt 2.3 wurde bereits verwendet, daß Verteilungen, die auf einer beschränkten Trägermenge definiert sind, also etwa die Betaverteilungen oder die Verteilung von Q , durch ihre Momente $E(X^n)$, $n \in \mathbb{N}$ eindeutig bestimmt werden. Aber auch für die auf dem halboffenen Träger $\mathbb{R}^+$ definierte Familie der Gammaverteilungen gilt diese Eigenschaft, was sich mit Hilfe des folgenden, bei Kendall, Stuart und Ord (1987, S. 137) angegebenen hinreichenden Kriteriums überprüfen läßt. Gilt

$$\limsup_{n \rightarrow \infty} \frac{\sqrt[n]{E(|X|^n)}}{n} < \infty,$$

so bestimmen die Momente $E(X^n)$, $n \in \mathbb{N}$ eindeutig die Verteilung von X . Für die Familie der Gammaverteilungen bleibt zu zeigen, daß

$$\limsup_{n \rightarrow \infty} \frac{\sqrt[n]{\frac{\Gamma(a+n)}{\Gamma(a)} b^n}}{n} < \infty$$

erfüllt ist. Aus $\frac{\Gamma(a+n)}{\Gamma(a)} = (a+n-1) \cdot \dots \cdot a \leq (a+n-1)^n$ folgt aber mit

$$\limsup_{n \rightarrow \infty} \frac{\sqrt[n]{E(|X|^n)}}{n} \leq b$$

sofort die Behauptung.

Die Momente von $\ln Q$

Aus den Momenten der Zufallsvariablen Q können in einfacher Weise die Momente von $\ln Q$ bestimmt werden. Die für $r > -an$ existierenden Momente

$$\mu(r) \stackrel{\text{def}}{=} E(Q^r) = \frac{\Gamma(an) \Gamma(a + \frac{r}{n})^n}{\Gamma(a)^n \Gamma(an+r)} n^r$$

definieren über die Beziehung

$$\mu(r) = E\left(e^{-r(-\ln Q)}\right)$$

die Laplace-Transformierte der positiven Zufallsvariablen $-\ln Q$. Damit läßt sich für $k \in \mathbb{N}$ das Moment $E((\ln Q)^k)$ durch k -faches differenzieren von μ nach r und den Grenzübergang $r \rightarrow 0$ bilden (s. Feller (1971, S. 435)):

$$E((\ln Q)^k) = \lim_{r \rightarrow 0} \frac{\partial^k}{\partial r^k} \mu(r). \quad (2.38)$$

Die Ableitung von μ lautet

$$\begin{aligned} \frac{\partial}{\partial r} \mu(r) &= \mu(r) \left[\psi\left(a + \frac{r}{n}\right) - \psi(an+r) + \ln n \right] \\ &= \mu(r) \left[\rho\left(a + \frac{r}{n}\right) - \rho(an+r) \right], \end{aligned}$$

und aus $\mu(0) = 1$ folgt damit

$$E(\ln Q) = \rho(a) - \rho(an). \quad (2.39)$$

Die k -te Ableitung $\mu^{(k)}$ von μ läßt sich unter Einführung der Funktion g

$$g(r) \stackrel{\text{def}}{=} \psi\left(a + \frac{r}{n}\right) - \psi(an + r) + \ln n \quad (2.40)$$

unter Anwendung der Leibnizschen Produktregel (s. Bronstein, Semendjajew (1987, S. 269)) in der folgenden Weise darstellen:

$$\mu^{(k)}(r) = \sum_{i=0}^{k-1} \binom{k-1}{i} \mu^{(i)}(r) g^{(k-1-i)}(r), \quad k \geq 1,$$

wobei $\mu^{(0)}$ und $g^{(0)}$ die Funktionen μ und g bezeichnen. Die k -te Ableitung der Funktion g

$$g^{(k)}(r) = \frac{1}{n^k} \psi^{(k)}\left(a + \frac{r}{n}\right) - \psi^{(k)}(an + r) \quad (2.41)$$

läßt sich durch die ebenfalls bei Spanier, Oldham (1987, S. 429ff) sowie Abramowitz, Stegun (1984, S. 81ff) diskutierte Polygammafunktion $\psi^{(k)}$, beschreiben. Für $k \in \mathbb{N}$ existiert die Polygammafunktion ebenso wie die Digammafunktion ψ im Bereich $\mathbb{R}^+$, so daß die Funktion $g^{(k)}$ für $r > -an$ definiert ist, woraus schließlich die Existenz der Momente $E((\ln Q)^k)$, $k \in \mathbb{N}$ folgt:

$$E((\ln Q)^k) = \begin{cases} 1 & \text{für } k = 0 \\ \sum_{i=0}^{k-1} \binom{k-1}{i} E((\ln Q)^i) g^{(k-1-i)}(0) & \text{für } k \geq 1. \end{cases} \quad (2.42)$$

Die asymptotische Verteilung von $-2an \ln Q$

Aus (2.42) ergibt sich

$$E((-2an \ln Q)^k) = \begin{cases} 1 & \text{für } k = 0 \\ \sum_{i=0}^{k-1} \binom{k-1}{i} E((-2an \ln Q)^i) (-2an)^{k-i} g^{(k-1-i)}(0) & \text{für } k \geq 1. \end{cases} \quad (2.43)$$

Aus den in Abschnitt 2.4 vorgestellten asymptotischen Entwicklungen der Funktion ρ (2.29) und (2.30) folgt durch Einsetzen in (2.39)

$$E(-2an \ln Q) = -2an(\rho(a) - \rho(an)) = \begin{cases} 2(n-1) + o(1) & \text{für } a \rightarrow 0 \\ n-1 + o(1) & \text{für } a \rightarrow \infty. \end{cases} \quad (2.44)$$

Für die Polygammafunktion $\psi^{(k)}$ ergibt sich aus der Rekursionsbeziehung

$$\psi^{(k)}(x) = \psi^{(k)}(x+1) + (-1)^{k+1} \frac{k!}{x^{k+1}}, \quad k \geq 0$$

beim Grenzübergang $x \rightarrow 0$, mit $\lim_{x \rightarrow 0} |\psi^{(k)}(x+1)| = |\psi^{(k)}(1)| < \infty$

$$\psi^{(k)}(x) = (-1)^{k+1} \frac{k!}{x^{k+1}} + o\left(\frac{1}{x^{k+1}}\right) \quad \text{für } x \rightarrow 0$$

und damit

$$(-2an)^{k+1} g^{(k)}(0) = 2^{k+1} k! (n-1) + o(1) \quad \text{für } a \rightarrow 0.$$

Aus der für $k \geq 1$ gegebenen asymptotischen Reihe

$$\psi^{(k)}(x) \sim (-1)^{k+1} \left[\frac{(k-1)!}{x^k} + \frac{k!}{2x^{k+1}} + \sum_{j=1}^{\infty} B_{2j} \frac{(2j+k-1)!}{(2j)! x^{2j+k}} \right] \quad \text{für } x \rightarrow \infty$$

(B_{2j} : Bernoulli-Zahlen) oder kurz

$$\psi^{(k)}(x) = (-1)^{k+1} \frac{(k-1)!}{x^k} \left(1 + \frac{k}{2x} \right) + o\left(\frac{1}{x^{k+1}}\right) \quad \text{für } x \rightarrow \infty$$

folgt

$$(-2an)^{k+1} g^{(k)}(0) = 2^k k! (n-1) + o(1) \quad \text{für } a \rightarrow \infty.$$

Einsetzen in (2.43) ergibt für $k \geq 1$ die folgenden asymptotischen Momente:

$$\begin{aligned} E\left((-2an \ln Q)^k\right) \\ = \begin{cases} 2^k (n-1) (k-1)! \sum_{i=0}^{k-1} \frac{1}{2^i i!} \left[E\left((-2an \ln Q)^i\right) + o(1) \right] & \text{für } a \rightarrow 0 \\ 2^{k-1} (n-1) (k-1)! \sum_{i=0}^{k-1} \frac{1}{2^i i!} \left[E\left((-2an \ln Q)^i\right) + o(1) \right] & \text{für } a \rightarrow \infty. \end{cases} \end{aligned}$$

Durch Induktion lassen sich daraus die für $k \geq 2$ gültigen Rekursionsbeziehungen

$$E\left((-2an \ln Q)^k\right) = \begin{cases} 2(k+n-2) \left[E\left((-2an \ln Q)^{k-1}\right) + o(1) \right] & \text{für } a \rightarrow 0 \\ 2\left(k + \frac{n-3}{2}\right) \left[E\left((-2an \ln Q)^{k-1}\right) + o(1) \right] & \text{für } a \rightarrow \infty \end{cases}$$

herleiten, die unter Verwendung von (2.44) zum Ergebnis

$$E\left((-2an \ln Q)^k\right) = \begin{cases} 2^k \frac{\Gamma(n-1+k)}{\Gamma(n-1)} + o(1) & \text{für } a \rightarrow 0 \\ 2^k \frac{\Gamma\left(\frac{n-1}{2}+k\right)}{\Gamma\left(\frac{n-1}{2}\right)} + o(1) & \text{für } a \rightarrow \infty. \end{cases} \quad (2.45)$$

führen. Diese asymptotischen Momente sind die Momente der χ^2 -Verteilung mit $2(n-1)$ (für $a \rightarrow 0$) bzw. $n-1$ Freiheitsgraden (für $a \rightarrow \infty$). Nach einem bei Kendall, Stuart und Ord (1987, S. 144) angegebenen Satz folgt daraus auch die Verteilungskonvergenz der Verteilung von $-2an \ln Q$ gegen die entsprechende

χ^2 -Verteilung, da sämtliche Momente existieren, und die Grenzwerte der Momente eindeutig die Grenzverteilung bestimmen. Die Verteilungskonvergenz läßt sich auch einfacher über eine Konvergenzaussage der Laplace-Transformierten gegen die Laplace-Transformierte der χ^2 -Verteilung zeigen. Die Konvergenz und sogar die Existenz der Momente folgt daraus jedoch noch nicht. Bain und Engelhardt (1975) geben die Kumulantenerzeugendenfunktion der Variablen $-\ln Q$ an und berechnen daraus die ersten zwei Momente der Verteilung von $-2an \ln Q$. Weiter stellen sie die oben angegebenen Konvergenzen dieser Momente gegen die ersten beiden Momente der χ^2 -Verteilung fest.

2.6. Die asymptotische Verteilung von $\hat{a}_{ML}$

Im folgenden soll gezeigt werden, daß die folgenden Aussagen gelten:

$$\lim_{a \rightarrow 0} E \left(\left(\frac{2an}{\hat{a}_{ML}} \right)^r \right) = 2^r \frac{\Gamma(n-1+r)}{\Gamma(n-1)}, \quad r > -(n-1)$$

(2.46)

und

$$\lim_{a \rightarrow \infty} E \left(\left(\frac{an}{\hat{a}_{ML}} \right)^r \right) = 2^r \frac{\Gamma\left(\frac{n-1}{2} + r\right)}{\Gamma\left(\frac{n-1}{2}\right)}, \quad r > -\frac{n-1}{2},$$

woraus analog zu der entsprechenden Aussage für die Verteilung von $-2an \ln Q$ für $a \rightarrow 0$ bzw. $a \rightarrow \infty$ die Verteilungskonvergenz der Verteilungen von $\frac{2an}{\hat{a}_{ML}}$ bzw. $\frac{an}{\hat{a}_{ML}}$ gegen die χ^2 -Verteilungen mit $2(n-1)$ bzw. $n-1$ Freiheitsgraden folgt.

Vorbereitend hierzu werden die folgenden Konvergenzaussagen benötigt:

Q konvergiert für $a \rightarrow 0$ stochastisch gegen 0

und

Q konvergiert für $a \rightarrow \infty$ stochastisch gegen 1.

Aus der Umformung

$$E(Q) = \frac{\Gamma(an) \Gamma\left(a + \frac{1}{n}\right)^n}{\Gamma(a)^n \Gamma(an+1)} n = a^{n-1} \frac{\Gamma(an+1)}{\Gamma(a+1)^n} \frac{\Gamma\left(a + \frac{n+1}{n}\right)}{\left(a + \frac{1}{n}\right)^n \Gamma(an+1)}$$

folgt

$$\lim_{a \rightarrow 0} E(Q) = 0,$$

worauf die Anwendung der Tschebyschev-Markov Ungleichung für $q > 0$ mit

$$\lim_{a \rightarrow 0} P_a \{Q \geq q\} \leq \lim_{a \rightarrow 0} \frac{E(Q)}{q} = 0$$

die Behauptung liefert.

Einsetzen der Stirlingschen Formel

$$\ln \Gamma(x) = \left(x - \frac{1}{2}\right) \ln x - x + \frac{1}{2} \ln(2\pi) + o(1) \quad \text{für } x \rightarrow \infty$$

in $E(Q)$ ergibt

$$E(Q) = \left(\frac{a}{a + \frac{1}{n}}\right)^{\frac{n-1}{2}} + o(1) \quad \text{für } a \rightarrow \infty$$

und führt damit zum Grenzwert

$$\lim_{a \rightarrow \infty} E(Q) = 1.$$

Die Tschebyschev-Markov Ungleichung auf die Zufallsvariable $1 - Q$ angewendet liefert für $q > 0$

$$P_a \{1 - Q \geq q\} \leq \frac{1 - E(Q)}{q}$$

und daraus die Behauptung

$$\lim_{a \rightarrow \infty} P_a \{Q \leq 1 - q\} = 0.$$

Zwischen den Zufallsvariablen $\hat{a}_{ML}$ und Q wurde in Abschnitt 2.4 der Zusammenhang

$$\hat{a}_{ML} = \frac{\tau(\ln Q)}{\ln Q}$$

hergeleitet. Aus der dort gezeigten Grenzwertbeziehung (2.36) ergibt sich

$$1 + \frac{1}{\tau(\ln Q)} \text{ konvergiert für } a \rightarrow 0 \text{ stochastisch gegen } 0$$

und

$$1 + \frac{1}{2\tau(\ln Q)} \text{ konvergiert für } a \rightarrow \infty \text{ stochastisch gegen } 0.$$

Die Differenzen

$$-2an \ln Q - \frac{2an}{\hat{a}_{ML}} = -2an \ln Q \left(1 + \frac{1}{\tau(\ln Q)}\right)$$

und

$$-2an \ln Q - \frac{an}{\hat{a}_{ML}} = -2an \ln Q \left(1 + \frac{1}{2\tau(\ln Q)}\right)$$

stellen für $a \rightarrow 0$ bzw. $a \rightarrow \infty$ das Produkt einer χ^2 -verteilten mit einer stochastisch gegen 0 konvergenten Zufallsvariablen dar und konvergieren somit selbst stochastisch gegen 0, woraus die stochastische Konvergenz und damit

auch die Verteilungskonvergenz der Zufallsvariablen $\frac{2an}{\hat{a}_{ML}}$ bzw. $\frac{an}{\hat{a}_{ML}}$ gegen die χ^2 -Verteilungen mit $2(n-1)$ bzw. $n-1$ Freiheitsgraden für $a \rightarrow 0$ bzw. $a \rightarrow \infty$ folgt.

Aus der Abschätzung (2.32) ergibt sich

$$\frac{1}{2}(-2an \ln Q) \leq \frac{an}{\hat{a}_{ML}} \leq -2an \ln Q.$$

Da die Momente von $-2an \ln Q$ wie gezeigt existieren, folgt somit auch die Existenz der Momente von $\frac{2an}{\hat{a}_{ML}}$ und $\frac{an}{\hat{a}_{ML}}$, die zudem — da die Momente von $-2an \ln Q$ für $a \rightarrow 0$ und $a \rightarrow \infty$ konvergieren — bei diesen Grenzübergängen ebenfalls beschränkt bleiben. Die Verteilungskonvergenz sowie die Existenz und Beschränktheit der Momente von $\frac{an}{\hat{a}_{ML}}$ erlauben die Anwendung des zweiten Grenzwertsatzes (s. Kendall, Stuart und Ord (1987, S. 143)) und liefern damit die Konvergenz gegen die Momente der Grenzverteilung (Behauptung (2.46)).

Aus den Momenten der χ^2 -Verteilung kann nun beispielsweise

$$\begin{aligned} \lim_{a \rightarrow 0} E\left(\frac{\hat{a}_{ML}}{a}\right) &= \frac{n}{n-2}, \quad n > 2 \\ \lim_{a \rightarrow \infty} E\left(\frac{\hat{a}_{ML}}{a}\right) &= \frac{n}{n-3}, \quad n > 3 \end{aligned}$$

berechnet werden, was für kleine Stichprobenumfänge auf eine deutliche Verzerrung hinweist. Choi und Wette (1969) stellen im Rahmen einer Monte-Carlo Untersuchung die Vermutung auf, daß für alle $a > 0$

$$E\left(\frac{\hat{a}_{ML}}{a}\right) > 1$$

gilt, Berman (1981) beweist diese Behauptung. Bowman und Shenton (1988, S. 84ff) berechnen asymptotische Reihenentwicklungen der ersten vier Momente für $a \rightarrow \infty$. Für den Erwartungswert von $\hat{a}_{ML}$ lautet ihr Ergebnis beispielsweise

$$E(\hat{a}_{ML}) = \frac{n}{n-3}a - \frac{2}{3(n-3)} + \frac{n-1}{9(n-3)na} + o\left(\frac{1}{a}\right), \quad \text{für } a \rightarrow \infty.$$

(Die Autoren geben noch weitere Terme bis zur Ordnung $\frac{1}{a^4}$ an).

2.7. Die asymptotische Verteilung von $\hat{\sigma}_{ML}^2$

Aus der Unabhängigkeit von $\bar{X}$ und $\hat{a}_{ML}$ lassen sich nun auf einfache Weise Aussagen über die Verteilung des Maximum-Likelihood-Schätzers $\hat{\sigma}_{ML}^2 = \frac{\bar{X}^2}{\hat{a}_{ML}}$ für die Varianz treffen. Die Zufallsvariable $\bar{X}$ ist gammaverteilt mit Gestaltparameter an und Skalierungsparameter $\frac{b}{n} = \frac{\sigma}{n\sqrt{a}}$, so daß das $2r$ -te Moment für $2r > -an$

$$E(\bar{X}^{2r}) = \frac{\Gamma(an+2r)}{\Gamma(an)} \frac{\sigma^{2r}}{n^{2r}a^r} \quad (2.47)$$

lautet. Aus der Unabhängigkeit von $\tilde{X}$ und $\hat{a}_{ML}$ folgt

$$E(\hat{\sigma}_{ML}^{2r}) = E\left(\left(\frac{\tilde{X}^2}{an}\right)^r\right) E\left(\left(\frac{an}{\hat{a}_{ML}}\right)^r\right).$$

Durch Einsetzen von (2.46) und (2.47) ergibt sich unter Verwendung von

$$\lim_{a \rightarrow 0} \frac{\Gamma(an + 2r)}{an\Gamma(an)} = \Gamma(2r)$$

für $a \rightarrow 0$

$$\lim_{a \rightarrow 0} \frac{1}{an} E\left((an\hat{\sigma}_{ML})^{2r}\right) = \left(\frac{\sigma^2}{n}\right)^r \frac{\Gamma(n-1+r)}{\Gamma(n-1)} \Gamma(2r). \quad (2.48)$$

Für $r = 1$ ergibt sich daraus beispielsweise

$$\lim_{a \rightarrow 0} an E(\hat{\sigma}_{ML}^2) = \sigma^2 \frac{n-1}{n},$$

für $r = 2$

$$\lim_{a \rightarrow 0} (an)^3 E(\hat{\sigma}_{ML}^4) = 6\sigma^4 \frac{n-1}{n}.$$

Unter Anwendung der Stirlingschen Formel lautet der Grenzwert der Momente für $a \rightarrow \infty$

$$\lim_{a \rightarrow \infty} E(\hat{\sigma}_{ML}^{2r}) = \left(2\frac{\sigma^2}{n}\right)^r \frac{\Gamma\left(\frac{n-1}{2} + r\right)}{\Gamma\left(\frac{n-1}{2}\right)}.$$

Dies sind die Momente der Gammaverteilung mit Gestaltparameter $\frac{n-1}{2}$ und Skalierungsparameter $2\frac{\sigma^2}{n}$, so daß sich als Ergebnis für $a \rightarrow \infty$ die Verteilungskonvergenz der Verteilung von

$$\frac{n\hat{\sigma}_{ML}^2}{\sigma^2}$$

gegen die χ^2 -Verteilung mit $n-1$ Freiheitsgraden ergibt. Diese Aussage ist nicht überraschend. Der Grenzübergang $a \rightarrow \infty$ bedeutet schließlich — salopp gesprochen — eine immer bessere Annäherung der zugrundeliegenden Gammaverteilung an eine Normalverteilung, so daß die Verteilungskonvergenz gegen die Verteilung des Maximum-Likelihood-Schätzers der Varianz in der Normalverteilungssituation plausibel erscheint, auch wenn dieser völlig anders berechnet wird.

2.8. Einige Eigenschaften der Verteilungsfunktion von $\hat{a}_{ML}$

Die Verteilungsfunktion der Zufallsvariablen $\hat{a}_{ML}$ soll im folgenden durch das Symbol $\mathcal{A}_{a,n}$ bezeichnet werden:

$$\mathcal{A}_{a,n}(\hat{a}) = P_a \{ \hat{a}_{ML} \leq \hat{a} \}.$$

In der obigen Darstellung wird $\mathcal{A}_{a,n}$ bei festem Parameterwert a ($a > 0$) als Funktion von $\hat{a}$ aufgefaßt. Wird $\mathcal{A}_{a,n}$ aber als Funktion des Parameters a bei festem $\hat{a}$ betrachtet, so soll dieses in der Notation $\tilde{\mathcal{A}}_{\hat{a},n}(a)$ zum Ausdruck kommen. Diese Notation bietet insbesondere eine einfache Unterscheidungsmöglichkeit der in Kapitel 4 benötigten Umkehrfunktionen $\mathcal{A}_{a,n}^{-1}$ und $\tilde{\mathcal{A}}_{\hat{a},n}^{-1}$. Wie in Abschnitt 2.3 soll stets $n \geq 2$ vorausgesetzt werden. Wie gezeigt, besitzt Q für $n = 1$ eine Einpunktverteilung an der Stelle 1, was aufgrund der Eigenschaften der Funktion ρ^{-1} bedeutet, daß $\hat{a}_{ML}$ stets den Wert ∞ annimmt. Für die Funktion $\mathcal{A}_{a,n}$ sind für $a > 0$ die folgenden Eigenschaften offenkundig erfüllt:

$\mathcal{A}_{a,n}$ ist stetig und streng isoton im Bereich $\mathbb{R}^+$ mit den Grenzwerten $\mathcal{A}_{a,n}(0) = 0$ und $\lim_{\hat{a} \rightarrow \infty} \mathcal{A}_{a,n}(\hat{a}) = 1$.

In folgenden sollen analoge Aussagen für die Funktion $\tilde{\mathcal{A}}_{\hat{a},n}$ hergeleitet werden. Es wird sich zeigen, daß für alle positiven $\hat{a}$ $\tilde{\mathcal{A}}_{\hat{a},n}$ eine im Bereich $\mathbb{R}^+$ streng antitone Funktion mit den Grenzwerten

$$\lim_{a \rightarrow 0} \tilde{\mathcal{A}}_{\hat{a},n}(a) = 1 \quad \text{und} \quad \lim_{a \rightarrow \infty} \tilde{\mathcal{A}}_{\hat{a},n}(a) = 0$$

ist. Für $\hat{a} \leq 0$ gilt trivialerweise

$$\forall a > 0 : \quad \tilde{\mathcal{A}}_{\hat{a},n}(a) \equiv 0,$$

so daß auch dieser Fall, ebenso wie der Fall $n = 1$ nicht mehr betrachtet wird.

Grenzwertaussagen

In Abschnitt 2.6 wurde die stochastische Konvergenz von Q gegen 0 für $a \rightarrow 0$ sowie gegen 1 für $a \rightarrow \infty$ gezeigt. Aus den in Abschnitt 2.4 gezeigten Grenzwertaussagen (2.26) zur Funktion ρ ergibt sich damit aufgrund der durch die Formel (2.27) ausgedrückten Beziehung zwischen den Verteilungsfunktionen von Q und $\hat{a}_{ML}$ die stochastische Konvergenz von $\hat{a}_{ML}$ gegen 0 für $a \rightarrow 0$ sowie die stochastische Konvergenz gegen ∞ für $a \rightarrow \infty$. Daraus folgen unmittelbar die obige Grenzwertaussagen.

Monotonie

Aus der Dichtefunktion von $\hat{a}_{ML}$ (s. 2.37)

$$f_{\hat{a}_{ML}}^{\hat{a}}(\hat{a}) = \frac{\Gamma(an)}{n^{an-1}\Gamma(a)^n} e^{an\rho(\hat{a})} \rho'(\hat{a}) \tilde{p}(e^{\rho(\hat{a})})$$

wird für festes $a > 0$ der für alle $\epsilon > -a$ definierte Dichtequotient

$$q_{\epsilon}(\hat{a}) \stackrel{\text{def}}{=} \frac{f_{a+\epsilon}^{\hat{a}ML}(\hat{a})}{f_a^{\hat{a}ML}(\hat{a})}$$

gebildet. Ableiten nach $\hat{a}$ liefert unmittelbar

$$\frac{\partial}{\partial \hat{a}} q_{\epsilon}(\hat{a}) = \epsilon n \rho'(\hat{a}) q_{\epsilon}(\hat{a}).$$

Der Dichtequotient besitzt damit aufgrund der strengen Isotonie der Funktion ρ für alle $\hat{a} > 0$ das gleiche Vorzeichen wie ϵ , so daß aufgrund der Normierungsbedingung

$$\forall a > 0 \quad \forall \epsilon > -a : \quad \int_0^{\infty} f_{a+\epsilon}^{\hat{a}ML}(\hat{a}) d\hat{a} = \int_0^{\infty} f_a^{\hat{a}ML}(\hat{a}) d\hat{a}$$

für alle positiven a_1 und a_2 mit $a_2 > a_1$ ein Wert $\hat{a}_0$ existiert, für den

$$f_{a_2}^{\hat{a}ML}(\hat{a}) < f_{a_1}^{\hat{a}ML}(\hat{a}) \iff \hat{a} < \hat{a}_0$$

erfüllt ist. Für $\hat{a} \leq \hat{a}_0$ folgt damit direkt die strenge Antitonie von $\tilde{\mathcal{A}}_{\hat{a},n}$, für $\hat{a} > \hat{a}_0$ folgt sie sofort aus der Ungleichung $1 - \int_{\hat{a}}^{\infty} f_{a_2}^{\hat{a}ML}(y) dy > 1 - \int_{\hat{a}}^{\infty} f_{a_1}^{\hat{a}ML}(y) dy$. Die Stetigkeit von $\tilde{\mathcal{A}}_{\hat{a},n}$ ergibt sich schließlich aus der Stetigkeit der Dichtefunktion (als Funktion des Parameters a aufgefaßt).

2.9. Einige Eigenschaften der Verteilungsfunktion $\mathcal{G}_{a,\sigma^2}$

Die Dichtefunktion einer mit Varianz σ^2 gammaverteilten Zufallsvariablen X lautet

$$f_{a,\sigma^2}(x) = \frac{(\sqrt{a})^a}{\Gamma(a) (\sqrt{\sigma^2})^a} x^{a-1} e^{-x\sqrt{\frac{a}{\sigma^2}}}.$$

Es soll gezeigt werden, daß für festes $x > 0$ und $\sigma^2 > 0$ die Verteilungsfunktion $\mathcal{G}_{a,\sigma^2}$ die Grenzwerte

$$\lim_{a \rightarrow 0} \mathcal{G}_{a,\sigma^2}(x) = 1 \quad \text{und} \quad \lim_{a \rightarrow \infty} \mathcal{G}_{a,\sigma^2}(x) = 0$$

besitzt. Dieses geschieht wieder mit Hilfe der Tschebyschev-Markov Ungleichung und den für festes σ^2 gültigen asymptotischen Erwartungswerten

$$\lim_{a \rightarrow 0} E(X) = \lim_{a \rightarrow 0} \frac{\Gamma(a+1)}{\Gamma(a)} \sqrt{\frac{\sigma^2}{a}} = 0$$

und

$$\lim_{a \rightarrow \infty} E(X^{-1}) = \lim_{a \rightarrow \infty} \frac{\Gamma(a-1)}{\Gamma(a)} \sqrt{\frac{a}{\sigma^2}} = 0.$$

Es gilt damit für alle $x > 0$

$$\lim_{a \rightarrow 0} P_{a,\sigma^2} \{X \geq x\} = 0 = \lim_{a \rightarrow \infty} P_{a,\sigma^2} \{X^{-1} \geq x\}$$

und damit die obige Behauptung.

Im Gegensatz zu $\tilde{\mathcal{A}}_{a,n}$ ist $\mathcal{G}_{a,\sigma^2}$ allerdings in a nicht monoton. Auf den Beweis dieser Behauptung soll an dieser Stelle verzichtet werden, er kann wiederum über die Betrachtung des Dichtequotienten geführt werden.

3. Die Berechnung der Verteilungsfunktion der Zufallsvariablen Q

Im letzten Kapitel wurden die Zufallsvariablen Q und $\hat{a}_{ML}$ eingeführt und diskutiert. Zur Bestimmung eines Konfidenzintervalls ist es natürlich nötig, ihre Verteilungen numerisch bestimmen zu können. Es wird sich im nächsten Kapitel zeigen, daß es nicht ausreicht, nur gewisse Quantile der Verteilungen berechnen zu können, sondern daß es notwendig ist, die Verteilungen an allen Stellen und für einen sehr großen Bereich des Parameters a mit möglichst großer Genauigkeit zu bestimmen. Da die Berechnung der Funktion ρ und ihrer Inversen ρ^{-1} aus numerischer Sicht kein Problem darstellt (s. z.B. Spanier, Oldham (1987, S. 427), Bowman, Shenton (1988, S. 31)), ist es aus dem Zusammenhang (2.27) leicht möglich, die beiden Verteilungsfunktionen auseinander herzuleiten.

In der Literatur gibt es einige Ansätze zur Berechnung der Verteilungsfunktionen von Q bzw. $\hat{a}_{ML}$. Neben dem bereits erwähnten Ansatz von Bishop und Nair (1939), existieren noch einige Näherungsverfahren für die Dichtefunktion von Q , aus denen durch Integration die Verteilungsfunktion berechnet werden kann (s. z.B. Hartley (1940) und Box (1949)). Ein weiterer Ansatz wird von Bowman und Shenton verfolgt (1988, S. 47ff). Die Autoren bestimmen die ersten vier — wie gezeigt exakt bestimmbaren — Momente der Verteilung von Q , bzw. näherungsweise die ersten vier Momente der Verteilung von $\hat{a}_{ML}$ und empfehlen, diese Momente in ein Standardverteilungssystem (z.B. die Systeme der Pearson-, oder Johnson-Kurven) einzusetzen. Solch eine Vorgangsweise kann sicherlich nur eine grobe erste Näherung der Verteilung darstellen. Zur Berechnung der Momente von $\hat{a}_{ML}$ verwenden Bowman und Shenton unter anderem auch eine Entwicklung der Dichtefunktion von $\ln Q$ nach Laguerre-Polynomen (S. 77ff), ein Ansatz der in Abschnitt 3.1.2 zu einer direkten Berechnung der Verteilungsfunktion von $\ln Q$ verwendet wird. Schließlich sei noch auf die ebenfalls bereits erwähnte Arbeit von Bain und Engelhardt (1975) verwiesen, in der die ersten beiden Momente der Verteilung von $-2an \ln Q$ zur Bestimmung der beiden Momente einer approximierenden Gammaverteilungsfunktion verwendet werden. Diese für die Erfordernisse dieser Arbeit im allgemeinen noch zu grobe Approximation entspricht exakt dem in Abschnitt 3.1.2 unter (3.12) und (3.13) vorgeschlagenen ersten Approximationsschritt zur Berechnung der Verteilungsfunktion von $\ln Q$.

3.1. Die Berechnung mit Hilfe orthogonaler Polynome

Im folgenden sollen zwei Ansätze zur Berechnung der Verteilungsfunktion von Q mit Hilfe orthogonaler Polynome vorgestellt werden. Diesen Ansätzen liegen die in den Kapiteln 2.5 und 2.3 hergeleiteten Verteilungskonvergenzen (2.45) und (2.21) zugrunde. Diese legen es nahe zu versuchen, die Verteilungsfunktion von $-2an \ln Q$ durch die Verteilungsfunktion einer χ^2 -Verteilung bzw. die Verteilungsfunktion von Q^{an} mit Hilfe einer Betaverteilung zu approximieren. Die Approximationsgüte wird dabei im ersten Fall für kleine und große Werte des Parameters a und im zweiten Fall für große Werte von a sowie kleine Stichprobenumfänge hoch sein. Die Approximation kann dabei noch beträchtlich verbessert werden, indem die

approximierende Verteilung so modifiziert wird, daß möglichst viele Momente mit den für alle a und n exakt berechenbaren Momenten der Verteilungen von $-2an \ln Q$ und Q^{an} übereinstimmen. Im ersten Abschnitt soll kurz die allgemeine Methodik, die auf beide Approximationen angewendet werden kann, vorgestellt werden.

3.1.1. Die allgemeine Methodik

Sei Z eine stetige Zufallsvariable mit Dichtefunktion f_Z und für ein gegebenes $k \in \mathbb{N}$ existenten und berechenbaren Momenten

$$\mu_j \stackrel{\text{def}}{=} E(Z^j), \quad j = 1, \dots, k. \quad (3.1)$$

Ist zudem bekannt, daß sich die Verteilung von Z durch eine Verteilung mit Dichtefunktion w „gut“ approximieren läßt, wobei die approximierende Verteilung einer bestimmten Klasse von Verteilungsfamilien angehört, so läßt sich eine allgemeine Methodik zur Verbesserung dieser Approximation angeben. Unter den Verteilungsfamilien, denen die approximierende Verteilung entstammen kann, befinden sich neben den hier betrachteten Gamma- und Betaverteilungen auch die Normalverteilungen — der zentrale Grenzwertsatz liefert hier im allgemeinen die Rechtfertigung für die Approximation. Zur Verbesserung der Approximation von f_Z wird w mit einem Polynom vom Grad k multipliziert

$$\hat{f}_Z(z) \stackrel{\text{def}}{=} w(z) \sum_{j=0}^k e_j z^j, \quad (3.2)$$

wobei die reellen Konstanten $e_0, \dots, e_k$ so gewählt werden, daß das Gleichungssystem

$$\int_{-\infty}^{\infty} z^j f_Z(z) dz = \int_{-\infty}^{\infty} z^j \hat{f}_Z(z) dz, \quad j = 0, \dots, k \quad (3.3)$$

erfüllt wird. Die Gleichung für $j = 0$ stellt die Normierungsbedingung dar, die restlichen Gleichungen „erzwingen“ die Gleichheit der ersten k Momente der approximierenden und der zu approximierenden Verteilung. Die Bestimmung der Konstanten $e_0, \dots, e_k$ geschieht, wie nachfolgend beschrieben wird, mit Hilfe eines Systems orthogonaler Polynome

$$p_m(z) \stackrel{\text{def}}{=} \sum_{j=0}^m c_{mj} z^j, \quad m = 0, \dots, k, \quad (3.4)$$

dessen Konstanten $(c_{mj})_{m=0, \dots, k, j=0, \dots, m}$ so gegeben sind, daß die Orthogonalitätsbeziehung

$$\int_{-\infty}^{\infty} w(z) p_m(z) p_n(z) dz = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}, \quad m, n \in \{0, \dots, k\} \quad (3.5)$$

erfüllt ist. Ist w eine Normal-, Gamma- oder Betaverteilungsdichte, so gewährleisten die Systeme der Hermite-, verallgemeinerten Laguerre- oder Jacobipolynome die Erfüllung dieser Orthogonalitätsbeziehung. Eine umfangreiche Aufstellung einer Liste von Funktionen w und dazugehörigen Polynomsystemen ist z.B. bei Abramowitz, Stegun (1984, S. 332ff) zu finden, wobei dort allerdings auf die in (3.5) angegebene Normierung durch die Konstante 1 sowie die Normierung der Funktion w verzichtet wird, so daß die dort angegebenen Koeffizienten in elementarer Weise unnormiert werden müssen.

Um die gesuchten Konstanten $e_0, \dots, e_k$ zu bestimmen, wird das Polynom $\sum_{j=0}^k e_j z^j$ als Linearkombination der Polynome $p_0, \dots, p_k$ dargestellt:

$$\hat{f}_Z(z) = w(z) \sum_{m=0}^k d_m p_m(z), \quad (3.6)$$

woraus sich nach Bestimmung der Konstanten $d_0, \dots, d_k$

$$e_j = \sum_{m=j}^k d_m c_{mj}, \quad j = 0, \dots, k \quad (3.7)$$

ergibt. Der Vektor $d_0, \dots, d_k$ kann aus den Momenten $\mu_0, \dots, \mu_k$ und dem System $(c_{mj})_{m=0, \dots, k, j=0, \dots, m}$ berechnet werden, indem die beiden Seiten der Gleichung (3.6) für $n \in \{0, \dots, k\}$ mit dem Polynom p_n multipliziert und anschließend integriert werden:

$$\begin{aligned} \int_{-\infty}^{\infty} p_n(z) \hat{f}_Z(z) dz &= \int_{-\infty}^{\infty} w(z) p_n(z) \sum_{m=0}^k d_m p_m(z) dz \\ \Leftrightarrow \sum_{j=0}^n c_{nj} \int_{-\infty}^{\infty} z^j \hat{f}_Z(z) dz &= \sum_{m=0}^k d_m \int_{-\infty}^{\infty} p_n(z) p_m(z) w(z) dz. \end{aligned}$$

Aus der Momentenforderung (3.3) und der Orthogonalitätsbeziehung (3.5) folgt nun unmittelbar

$$d_n = \sum_{j=0}^n c_{nj} \mu_j, \quad n = 0, \dots, k. \quad (3.8)$$

Die gesuchten Konstanten $e_0, \dots, e_k$ ergeben sich schließlich durch Einsetzen von $d_0, \dots, d_k$ in (3.7):

$$e_j = \sum_{m=j}^k c_{mj} \sum_{l=0}^m c_{ml} \mu_l, \quad j = 0, \dots, k. \quad (3.9)$$

Die Approximation der Verteilungsfunktion

Die approximative Verteilungsfunktion

$$\hat{F}_Z(z) = \int_{-\infty}^z \hat{f}_Z(x) dx$$

wird durch Einsetzen von (3.2) bestimmt:

$$\hat{F}_Z(z) = \sum_{j=0}^k e_j \int_{-\infty}^z x^j w(x) dx. \quad (3.10)$$

Es bleiben also noch die Integrale

$$I_j(z) \stackrel{\text{def}}{=} \int_{-\infty}^z x^j w(x) dx, \quad j = 0, \dots, k \quad (3.11)$$

zu berechnen. Im Falle der Gamma- und Betaverteilungen wird in den Abschnitten 3.1.2 und 3.1.3 je eine rekursive Beziehung angegeben, die für $j \geq 1$ einen Zusammenhang zwischen $I_j(z)$ und $I_{j-1}(z)$ herstellt, so daß nur noch das Integral $I_0(z) = \int_{-\infty}^z w(x) dx$ numerisch zu bestimmen bleibt. Die Berechnung der Verteilungsfunktion der Gamma- und Beta-Verteilung ist aber ein Standardproblem der numerischen Mathematik, für welches es sehr gute Verfahren gibt.

Zur Wahl von k

Die Berechnung der Approximation kann auf zweierlei Weise geschehen. Einerseits kann ein fester Wert von k vorgegeben werden, andererseits kann die Folge der Approximationen $\hat{F}_Z^{(k)}(z)$, $k = 0, 1, 2, \dots$ betrachtet werden. Die Vorgabe eines festen Wertes von k ist dabei nicht unproblematisch, da ein zu kleiner Wert eine noch zu ungenaue Approximation bedeutet, während für zu große k beträchtliche numerische Fehler auftreten können. Aus diesem Grunde ist es sinnvoll den zweiten Weg einzuschlagen, wodurch sich die Notwendigkeit ergibt, die Darstellung so zu wählen, daß es mit geringem Rechenaufwand möglich ist, aus $\hat{F}_Z^{(k)}(z)$ die Näherung $\hat{F}_Z^{(k+1)}(z)$ zu bestimmen.

3.1.2. Die Entwicklung nach verallgemeinerten Laguerre-Polynomen

In Kapitel 2.5 wurde gezeigt, daß $-2a \ln Q$ für $a \rightarrow 0$ nach Verteilung gegen eine χ^2 -Verteilung mit $2(n-1)$ bzw. für $a \rightarrow \infty$ gegen eine χ^2 -Verteilung mit $n-1$ Freiheitsgraden konvergiert.

Zur Wahl von w

Als Funktion w bietet sich die Dichtefunktion einer Gammaverteilung an, deren Parameter γ (Gestaltparameter) und δ (Skalierungsparameter) so gewählt werden, daß die ersten beiden in (2.43) angegebenen Momente von $-2an \ln Q$ mit den ersten beiden Momenten der $G(\gamma, \delta)$ -Verteilung übereinstimmen. Selbstverständlich können γ und δ auch andere Werte zugewiesen bekommen, doch ist diese naheliegende Vorbesetzung aus numerischen Gründen sinnvoll. Eine andere Möglichkeit ist es, sich auf die Familie der χ^2 -Verteilungen zu beschränken — also $\delta = 2$ und $2\gamma \in \mathbb{N}$ zu wählen — um Formeln zur Berechnung der χ^2 -Verteilungsfunktion anwenden zu können, die sich nicht auf den Fall $2\gamma \notin \mathbb{N}$ verallgemeinern lassen (einige solcher Formeln finden sich bei Abramowitz, Stegun (1984, S. 416ff)).

Wird die oben beschriebene Anpassung der ersten beiden Momente durchgeführt, so ergeben sich die folgenden Bestimmungsgleichungen:

$$\begin{aligned}\gamma\delta &= E(-2an \ln Q) \\ \gamma\delta^2 &= \text{Var}(-2an \ln Q)\end{aligned}$$

und daraus

$$\begin{aligned}\gamma &= \frac{[E(-2an \ln Q)]^2}{\text{Var}(-2an \ln Q)} = \frac{[E(\ln Q)]^2}{\text{Var}(\ln Q)} \\ \delta &= \frac{\text{Var}(-2an \ln Q)}{E(-2an \ln Q)} = -2an \frac{\text{Var}(\ln Q)}{E(\ln Q)}.\end{aligned}\tag{3.12}$$

Die Zufallsvariable

$$Z \stackrel{\text{def}}{=} \frac{-2an \ln Q}{\delta} = \ln Q \frac{E(\ln Q)}{\text{Var}(\ln Q)}\tag{3.13}$$

besitzt damit eine Verteilung, deren erste beiden Momente mit den Momenten der Gammaverteilung mit Gestaltparameter γ und Skalierungsparameter 1 übereinstimmen.

Die Approximation der Dichtefunktion

Das System der im Sinne der Beziehung (3.5) zur Dichtefunktion w ($z > 0$)

$$w(z) = \frac{1}{\Gamma(\gamma)} z^{\gamma-1} e^{-z}$$

orthogonalen Polynome, ist das System der verallgemeinerten Laguerre-Polynome und lautet

$$p_m(z) = \sum_{j=0}^m c_{mj} z^j, \quad m = 0, \dots, k$$

mit

$$c_{mj} = (-1)^j \binom{m}{j} \frac{\sqrt{\Gamma(m+\gamma)\Gamma(\gamma)} m!}{m! \Gamma(\gamma+j)}, \quad m = 0, \dots, k, \quad j = 0, \dots, m \quad (3.14)$$

(s. Abramowitz, Stegun (1984, S. 332ff)). Durch Einsetzen in (3.8) ergibt sich der Vektor $d_0, \dots, d_k$ und damit gemäß (3.6) für $z > 0$ die Approximation

$$\hat{f}_Z(z) = w(z) \sum_{m=0}^k \frac{\Gamma(m+\gamma)\Gamma(\gamma)}{m!} \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \mu_j \right] \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} z^j \right]. \quad (3.15)$$

Auf die Angabe der Koeffizienten $e_0, \dots, e_k$ soll hier verzichtet werden, da die Berechnung der Approximation der Verteilungsfunktion von Z über die Formel (3.15) leichter ist.

Die Approximation der Verteilungsfunktion

Die Integration von $\hat{f}_Z$ liefert

$$\begin{aligned} \hat{F}_Z(z) \\ = \sum_{m=0}^k \frac{\Gamma(m+\gamma)\Gamma(\gamma)}{m!} \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \mu_j \right] \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \int_0^z x^j w(x) dx \right]. \end{aligned} \quad (3.16)$$

Das Integral $I_j(z) = \int_0^z x^j w(x) dx$, $j = 0, \dots, k$ ist im wesentlichen ein Gamma-verteilungsintegral:

$$\begin{aligned} I_j(z) &= \int_0^z \frac{1}{\Gamma(\gamma)} x^{\gamma+j-1} e^{-x} dx \\ &= \frac{\Gamma(\gamma+j)}{\Gamma(\gamma)} \int_0^z \frac{1}{\Gamma(\gamma+j)} x^{\gamma+j-1} e^{-x} dx. \end{aligned}$$

Entsprechend der in Kapitel 2.1 eingeführten Schreibweise für die Verteilungsfunktion der Gammaverteilung, gilt also für $j = 0, \dots, k$

$$I_j(z) = \frac{\Gamma(\gamma+j)}{\Gamma(\gamma)} \mathcal{G}_{\gamma+j, \gamma+j}(z).$$

Aus der mehrfachen Anwendung der Rekursionsbeziehung

$$\mathcal{G}_{\gamma+1, \gamma+1}(z) = \mathcal{G}_{\gamma, \gamma}(z) - \frac{z^\gamma e^{-z}}{\Gamma(\gamma+1)} \quad (3.17)$$

(s. Abramowitz, Stegun (1984, S. 82)) ergibt sich

$$I_j(z) = \frac{\Gamma(\gamma+j)}{\Gamma(\gamma)} \mathcal{G}_{\gamma,\gamma}(z) - \frac{\Gamma(\gamma+j)}{\Gamma(\gamma)} \sum_{l=0}^{j-1} \frac{z^{\gamma+l} e^{-z}}{\Gamma(\gamma+l+1)}. \quad (3.18)$$

Durch Einsetzen in (3.16) läßt sich die Darstellung

$$\begin{aligned} \hat{F}_Z(z) &= \mathcal{G}_{\gamma,\gamma}(z) \sum_{m=0}^k \frac{\Gamma(m+\gamma)}{m!} \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \mu_j \right] \left[\sum_{j=0}^m (-1)^j \binom{m}{j} \right] \\ &\quad - \sum_{m=1}^k \frac{\Gamma(m+\gamma)}{m!} \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \mu_j \right] \left[\sum_{j=1}^m (-1)^j \binom{m}{j} \sum_{l=0}^{j-1} \frac{z^{\gamma+l} e^{-z}}{\Gamma(\gamma+l+1)} \right] \end{aligned}$$

herleiten, was sich unter Verwendung der aus dem binomischen Satz folgenden Beziehung

$$\sum_{j=0}^m (-1)^j \binom{m}{j} = \begin{cases} 0 & m \geq 1 \\ 1 & m = 0 \end{cases} \quad (3.19)$$

zu

$$\begin{aligned} \hat{F}_Z(z) &= \mathcal{G}_{\gamma,\gamma}(z) \\ &\quad - \sum_{m=1}^k \frac{\Gamma(m+\gamma)}{m!} \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \mu_j \right] \left[\sum_{j=1}^m (-1)^j \binom{m}{j} \sum_{l=0}^{j-1} \frac{z^{\gamma+l} e^{-z}}{\Gamma(\gamma+l+1)} \right] \end{aligned}$$

vereinfacht. Durch Umordnen und neuerliche Anwendung von (3.19) kann nun eine Darstellung der Form

$$\hat{F}_Z(z) = \mathcal{G}_{\gamma,\gamma}(z) + z^{\gamma} e^{-z} \sum_{l=0}^{k-1} \tau_{k,l} z^l \quad (3.20)$$

gewonnen werden. Die sich dabei ergebenden Koeffizienten $(\tau_{k,l})_{l=0, \dots, k-1}$ lauten

$$\tau_{k,l} = \frac{1}{\Gamma(\gamma+l+1)} \sum_{m=l+1}^k \frac{\Gamma(m+\gamma)}{m!} \left[\sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma+j)} \binom{m}{j} \mu_j \right] \left[\sum_{j=0}^l (-1)^j \binom{m}{j} \right]. \quad (3.21)$$

Die rekursive Berechnung von $\hat{F}_Z^{(k)}(z)$ aus $\hat{F}_Z^{(k-1)}(z)$

Für $k = 1$ ergibt sich

$$\tau_{1,0} = \frac{\gamma - \mu_1}{\Gamma(\gamma + 1)}.$$

Für $k > 1$ ist die folgende Rekursion nach dem Hornerschema (s. Bronstein, Semendjajew (1987, S. 780))

$$\begin{aligned} \hat{F}_Z^{(k)}(z) &= \hat{F}_Z^{(k-1)}(z) \\ &+ z^\gamma e^{-z} ((\dots((\lambda_{k,k-1}z + \lambda_{k,k-2})z + \lambda_{k,k-3})z + \dots + \lambda_{k,1})z + \lambda_{k,0}) \end{aligned} \quad (3.21)$$

anwendbar, wobei der Vektor $(\lambda_{k,l})_{l=0,\dots,k-1}$ durch

$$\lambda_{k,l} = \begin{cases} \tau_{k,l} - \tau_{k-1,l} & l = 0, \dots, k-2 \\ \tau_{k,k-1} & l = k-1 \end{cases}$$

gegeben ist. Mit

$$\eta_m \stackrel{\text{def}}{=} \sum_{j=0}^m \frac{(-1)^j}{\Gamma(\gamma + j)} \binom{m}{j} \mu_j$$

gilt

$$\tau_{k,k-1} = \frac{(-1)^{k+1}}{k!} \eta_k$$

und damit

$$\tau_{k,k-1} = -\frac{1}{k} \frac{\eta_k}{\eta_{k-1}} \tau_{k-1,k-2}.$$

Damit ist $\lambda_{k,k-1} = \tau_{k,k-1}$ bestimmt. Für $l < k-1$ lautet die Differenz $\lambda_{k,l} = \tau_{k,l} - \tau_{k-1,l}$

$$\lambda_{k,l} = \frac{1}{\Gamma(\gamma + l + 1)} \frac{\Gamma(\gamma + k)}{k!} \eta_k,$$

woraus sich zwischen $\lambda_{k,l}$ und $\lambda_{k,l+1}$ die Beziehung

$$\lambda_{k,l} = (\gamma + l + 1) \lambda_{k,l+1} + \frac{\Gamma(\gamma + k)}{\Gamma(\gamma + l + 1) k!} (-1)^l \binom{k}{l+1} \eta_k$$

ergibt. Schließlich kann auch der zweite Summand

$$\rho_{k,l} \stackrel{\text{def}}{=} \frac{\Gamma(\gamma + k)}{\Gamma(\gamma + l + 1) k!} (-1)^l \binom{k}{l+1} \eta_k, \quad l = 0, \dots, k-1$$

für $l < k-1$ rekursiv berechnet werden:

$$\rho_{k,l} = \begin{cases} -(\gamma + l + 1) \frac{l+2}{k-l-1} \rho_{k,l+1} & l = 0, \dots, k-2 \\ \lambda_{k,k-1} & l = k-1, \end{cases}$$

so daß die Koeffizienten $(\lambda_{k,l})_{l=k-1,\dots,1}$ durch das obige Schema in der richtigen Reihenfolge erzeugt werden.

Aus der Darstellung der Konstanten $\tau_{1,0}$ ist leicht ersichtlich, daß im Falle der zu Beginn dieses Abschnitts vorgeschlagenen Wahl der Parameter γ und δ $\hat{F}_Z^{(0)}(z)$ und $\hat{F}_Z^{(1)}(z)$ stets übereinstimmen, da $\tau_{1,0} = 0$ gilt. Desweiteren gilt sogar stets $\hat{F}_Z^{(0)}(z) = \hat{F}_Z^{(1)}(z) = \hat{F}_Z^{(2)}(z)$. Dieses ist plausibel, da die vorgeschlagene Wahl der Parameter ja gerade die Anpassung der ersten beiden Momente der approximierenden an die Momente der zu approximierenden Verteilung bedeutet.

3.1.3. Die Entwicklung nach Jacobi-Polynomen

In Kapitel 2.3 wurde unter (2.21) gezeigt, daß für $a \rightarrow \infty$ die Zufallsvariablen Q^{an} und $Y^{a(n-1)}$, wobei Y Beta $(a(n-1), \frac{n-1}{2})$ -verteilt sei, nach Verteilung gegeneinander konvergieren. Für $n = 2$ sind Q^2 und Y und damit Q^{2a} und Y^a sogar für alle $a > 0$ identisch verteilt.

Zur Wahl von w

Aufgrund dieses Sachverhaltes bietet es sich an,

$$Z \stackrel{\text{def}}{=} (Q^{an})^{\frac{1}{a(n-1)}} = Q^{\frac{n}{n-1}}$$

zu setzen und für w eine Betaverteilungsdichte ($0 < z < 1$)

$$w(z) = \frac{1}{B(p, q)} z^{p-1} (1-z)^{q-1}$$

zu wählen. Die Parameter p und q der Dichte können wiederum so bestimmt werden, daß die ersten beiden Momente der Verteilungen von Q^{an} und $Y^{a(n-1)}$ übereinstimmen. Dieses Vorgehen liefert

$$\begin{aligned} p &= \mu_1 \frac{\mu_1 - \mu_2}{\mu_2 - \mu_1^2} \\ q &= (1 - \mu_1) \frac{\mu_1 - \mu_2}{\mu_2 - \mu_1^2}, \end{aligned}$$

wobei zu beachten ist, daß entsprechend der Definition (3.1) mit μ_j das j -te Moment von Z bezeichnet wird, also $\mu_j = E(Q^j \frac{n}{n-1})$ gilt.

Die Approximation der Dichtefunktion

Die Koeffizienten des zur Dichtefunktion w orthogonalen Systems der Jacobi-Polynome (s. Abramowitz, Stegun (1984, S. 332ff))

$$p_m(z) = \sum_{j=0}^m c_{mj} z^j, \quad m = 0, \dots, k$$

lauten

$$c_{mj} = \sqrt{B(p, q)} (-1)^m \sqrt{\frac{\Gamma(p+m)(p+q+2m-1)}{m! \Gamma(q+m) \Gamma(p+q+m-1)}} \cdot (-1)^j \frac{1}{\Gamma(p+j)} \binom{m}{j} \Gamma(p+q+m+j-1), \quad m=0, \dots, k, \quad j=0, \dots, m. \quad (3.23)$$

Einsetzen dieser Koeffizienten in die Darstellung

$$\hat{f}_Z(z) = w(z) \sum_{j=0}^k \sum_{m=j}^k c_{mj} z^j \sum_{l=0}^m c_{ml} \mu_l$$

(s. (3.2) und (3.9)) liefert

$$\begin{aligned} \hat{f}_Z(z) = w(z) B(p, q) \sum_{m=0}^k \frac{1}{m!} \frac{\Gamma(p+m)(p+q+2m-1)}{\Gamma(q+m) \Gamma(p+q+m-1)} \\ \cdot \left[\sum_{j=0}^m (-1)^j \frac{\Gamma(p+q+m+j-1)}{\Gamma(p+j)} \binom{m}{j} \mu_j \right] \\ \cdot \left[\sum_{j=0}^m (-1)^j \frac{\Gamma(p+q+m+j-1)}{\Gamma(p+j)} \binom{m}{j} z^j \right]. \end{aligned} \quad (3.24)$$

Die Approximation der Verteilungsfunktion

Bei der Integration von $\hat{f}_Z$ ergeben sich wiederum die Integrale $I_j(z) = \int_0^z x^j w(x) dx$, $j = 0, \dots, k$. Durch Einsetzen der Dichtefunktion w ergibt sich

$$I_j(z) = \frac{B(p+j, q)}{B(p, q)} J_{p+j, q}(z),$$

wobei

$$J_{p+j, q}(z) \stackrel{\text{def}}{=} \int_0^z \frac{1}{B(p+j, q)} x^{p+j-1} (1-x)^{q-1} dx$$

die Verteilungsfunktion der Beta-Verteilung mit Parametern $p+j$ und q bezeichnet. Auch in diesem Fall existiert mit der Rekursionsbeziehung

$$J_{p+1, q}(z) = J_{p, q}(z) - z^p (1-z)^q \frac{\Gamma(p+q)}{\Gamma(p+1) \Gamma(q)}$$

(s. Abramowitz, Stegun (1984, S. 420)) die Möglichkeit, die Berechnung der Integrale $I_0(z), \dots, I_k(z)$ auf die Berechnung von $J_{p, q}(z)$ zurückzuführen. Durch

Integration der Approximation $\hat{f}_Z$ ergibt sich damit

$$\begin{aligned}\hat{F}_Z(z) &= J_{p,q}(z) \\ &\cdot \sum_{m=0}^k \frac{1}{m!} \frac{\Gamma(p+m)(p+q+2m-1)}{\Gamma(q+m)\Gamma(p+q+m-q)} \left[\sum_{j=0}^m (-1)^j \frac{\Gamma(p+q+m+j-1)}{\Gamma(p+j)} \binom{m}{j} \mu_j \right] \\ &\cdot \left[\sum_{j=0}^m (-1)^j \frac{\Gamma(p+q+m+j-1)\Gamma(q)}{\Gamma(p+q+j)} \binom{m}{j} \right] \\ &- \sum_{m=1}^k \frac{1}{m!} \frac{\Gamma(p+m)(p+q+2m-1)}{\Gamma(q+m)\Gamma(p+q+m-q)} \left[\sum_{j=0}^m (-1)^j \frac{\Gamma(p+q+m+j-1)}{\Gamma(p+j)} \binom{m}{j} \mu_j \right] \\ &\cdot \left[\sum_{j=1}^m (-1)^j \frac{\Gamma(p+q+m+j-1)\Gamma(q)}{\Gamma(p+q+j)} \binom{m}{j} \sum_{l=0}^{j-1} \frac{\Gamma(p+q+l)}{\Gamma(p+l+1)\Gamma(q)} z^{p+l}(1-z)^q \right].\end{aligned}$$

Da an der Stelle $z = 1$ für $q > 0$

$$z^{p+l}(1-z)^q \big|_{z=1} = 0$$

gilt und zudem aufgrund der Normierungsbedingung (3.3)

$$\hat{F}_Z(1) = J_{p,q}(1) = 1$$

erfüllt ist, folgt, daß der Faktor mit dem $J_{p,q}(z)$ multipliziert wird, den Wert 1 besitzt. Durch weitere Umformungen kann eine Darstellung der Gestalt

$$\hat{F}_Z(z) = J_{p,q}(z) + z^p(1-z)^q \sum_{l=0}^{k-1} \tau_{k,l} z^l$$

angegeben werden. Mit

$$\eta_m \stackrel{\text{def}}{=} \sum_{j=0}^m (-1)^j \frac{\Gamma(p+q+m+j-1)}{\Gamma(p+j)} \binom{m}{j} \mu_j, \quad m = 0, \dots, k \quad (3.25)$$

lauten die Koeffizienten

$$\begin{aligned}\tau_{k,l} &= \frac{\Gamma(p+q+l)}{\Gamma(p+l+1)} \sum_{m=l+1}^k \frac{1}{m!} \frac{\Gamma(p+m)(p+q+2m-1)}{\Gamma(q+m)\Gamma(p+q+m-1)} \eta_m \\ &\cdot \sum_{j=l+1}^m (-1)^{j+1} \frac{\Gamma(p+q+m+j-1)\Gamma(p+q+j)}{\Gamma(p+q+j)} \binom{m}{j}.\end{aligned} \quad (3.26)$$

In völliger Analogie zu der in Abschnitt 3.1.2 beschriebenen Vorgehensweise, kann auch in diesem Fall die Folge $\hat{F}_Z^{(0)}(z), \hat{F}_Z^{(1)}(z), \dots$ rekursiv berechnet werden. Auf die Angabe der entsprechenden Rekursionsbeziehungen zwischen den Größen $\tau_{k,l}, \rho_{k,l}$ und $\lambda_{k,l}$, die die gleiche Bedeutung wie bei der Rekursion des letzten Abschnitts besitzen, soll an dieser Stelle verzichtet werden, sie können aus (3.26) leicht hergeleitet werden.

3.2. Die Berechnung mit Hilfe einer Laplace-Transformation

Zu Beginn des Kapitels 2.5 wurde bereits erwähnt, daß der für $\lambda > -an$ definierte Erwartungswert

$$E(Q^\lambda) = E(e^{-\lambda(-\ln Q)})$$

die Laplace-Transformierte der Dichtefunktion der positiven Zufallsvariablen $-\ln Q$ an der Stelle λ darstellt. Die Laplace-Transformierte der Verteilungsfunktion von $-\ln Q$ ergibt sich durch Division der Laplace-Transformierten der Dichtefunktion durch λ , wie aus einem allgemeinen Zusammenhang folgt (s. Feller (1971, S. 435)). Der Wert der Laplace-Transformierten der Verteilungsfunktion von $-\ln Q$ an der Stelle λ lautet also

$$\frac{E(Q^\lambda)}{\lambda} = \frac{1}{\lambda} \frac{\Gamma(an) \Gamma(a + \frac{\lambda}{n})^n}{\Gamma(a)^n \Gamma(an + \lambda)} n^\lambda,$$

und ist für alle $\lambda > 0$ definiert. Damit kann ein Umkehrverfahren für die Laplace-Transformation angewendet werden, welches es erlaubt, aus der Laplace-Transformierten die Verteilungsfunktion von $-\ln Q$ zu berechnen. Ohne weiter auf die Theorie der Laplace-Transformation einzugehen, soll im folgenden eine bei Schremmer (1986, S. 30) hergeleitete Umkehrformel angewendet werden, die eine auf einfache Weise berechenbare und in allen Fällen — also für alle Werte von $a > 0$ und für alle $q \in (0, 1)$ — konvergente Reihenentwicklung der Verteilungsfunktion von $-\ln Q$ und aufgrund des Zusammenhangs

$$P_a\{Q \leq q\} = 1 - P_a\{-\ln Q \leq -\ln q\}$$

auch der Verteilungsfunktion von Q darstellt.

Die Umkehrformel

Sei $\tau > 0$ beliebig. Für die Verteilungsfunktion von Q gilt für alle $q \in (0, 1)$:

$$\forall a > 0: P_a\{Q \leq q\} = 1 - \sum_{m=1}^{\infty} c_m \sin((2m-1) \arccos(q^\tau)), \quad (3.27)$$

mit den Koeffizienten

$$c_m = \frac{\tau}{\pi} \sum_{k=1}^m (-1)^{m-k} 4^k \binom{m+k-2}{2k-2} \frac{E(Q^{(2k-1)\tau})}{(2k-1)\tau}, \quad m = 1, 2, \dots$$

Die Koeffizienten können in Erweiterung einer bei Schremmer (S. 28) angegebenen Beziehung dabei auf einfache Weise auch rekursiv bestimmt werden, wie sich durch einfache Umformungen bestätigen läßt:

$$c_1 = 4 \frac{E(Q^\tau)}{\pi}$$

$$c_m = 4^m \frac{E(Q^{(2m-1)\tau})}{(2m-1)\pi} - \sum_{k=1}^{m-1} e_{m,k} c_k, \quad m \geq 2$$

und

$$e_{m,m-1} = 2m - 3$$

$$e_{m,k} = e_{m,k+1} \frac{m+k}{m-k} \frac{2k-1}{2k+1}, \quad k = m-2, \dots, 1.$$

Zur Wahl der Konstanten τ

Die Konstante τ kann — wie oben festgestellt — beliebig aus dem Bereich der positiven reellen Zahlen gewählt werden, die Reihe konvergiert immer. Aus numerischer Sicht ist die Wahl von τ allerdings nicht so unproblematisch. Bei ungünstiger Wahl der Konstanten kann es sehr leicht zu sehr großen numerischen Fehlern kommen, eine auf die jeweilige Situation abgestimmte Wahl von τ ist unerlässlich. Als Ausgangspunkt für die dazu nötigen umfangreichen numerischen Vergleiche, können in der Arbeit von Schremmer (S. 61ff) vorgestellte Überlegungen dienen, die auf die vorliegende Fragestellung übertragen, den Vorschlag der Wahl von

$$\tau = \frac{1}{-\ln q} \frac{\ln 2}{2}$$

bedeuten. Diese Festsetzung erweist sich nicht nur aus numerischer Sicht als erfolgversprechend, sondern sie vereinfacht auch die Berechnungen, wie durch Einsetzen in (3.27) ersichtlich ist. Die Übertragung dieses aus einer wesentlich allgemeineren Betrachtungsweise gewonnenen Vorschlags auf die hier vorliegende spezielle Fragestellung gelingt recht gut, doch kann das numerische Verhalten des Verfahrens durch im allgemeinen relativ kleine Korrekturen dieses Vorschlags in gewissen Situationen noch deutlich verbessert werden. Diese Korrekturen hängen vom Wert des Parameters a , vom Wert des Arguments q sowie vom Stichprobenumfang n ab. Die Ergebnisse des nächsten Abschnitts werden zeigen, daß die numerischen Probleme, die bei der hier verwendeten Arithmetik auftreten, eine Anwendung dieses Verfahrens für mittlere und größere Stichprobenumfänge — und genau für diese ist die in Kapitel 4 besprochene Konfidenzintervallbildung überhaupt sinnvoll — nicht ratsam erscheinen lassen. Aus diesem Grunde soll auf eine detaillierte Beschreibung der Korrekturen verzichtet werden.

Zum Vergleich der drei Verfahren

Der grundsätzliche Vorteil dieses Verfahrens im Vergleich zu den beiden in Abschnitt 3.1 genannten Verfahren, besteht darin, daß einerseits Konvergenzaussagen vorliegen, andererseits das Verfahren mit sehr wenigen Annahmen auskommt (es werden nur die Momente $E(Q^{(2m-1)\tau})$, $m = 1, 2, \dots$ verwendet). Es ist damit prädestiniert, auch auf andere Situationen übertragen zu werden. Die Anwendbarkeit der beiden in Abschnitt 3.1 beschriebenen Verfahren beruht hingegen auf speziellen Zusatzinformationen. Während sich das numerische Verhalten des Umkehrverfahrens durch eine Ersetzung der hier verwendeten Arithmetik (64 Dualziffern entsprechend ca. 19–20 Dezimalstellen Genauigkeit) durch eine genauere im Prinzip beliebig steigern läßt, stehen numerische Probleme bei den in Abschnitt 3.1 genannten Verfahren nicht im Vordergrund, eine noch so genaue Arithmetik wird die Defizite, die zum Beispiel die Verteilungsfunktionsbestimmung

mit Hilfe von Jacobi-Polynomen für kleine Werte des Parameters a aufweist, nicht beheben können, hier fehlt eben die entsprechende theoretische Voraussetzung.

3.3. Einige numerische Beispiele

In den Tabellen 3.1 und 3.2 werden die drei in den Abschnitten 3.1 und 3.2 beschriebenen Verfahren hinsichtlich ihres numerischen Verhaltens verglichen.

Erläuterungen zu den Tabellen

In den Tabellen werden Vergleiche der Verfahren für die Stichprobenumfänge $n = 10, 20, 50, 250$ und 1000 sowie für die Parameterwerte $a = 0,1; 0,25; 0,5; 1; 2,5; 5; 7,5; 10; 25$ und 100 angegeben. Die Werte des Arguments q sind dabei jeweils so bestimmt, daß gilt

$$F_{Laguerre}^Q(q) = p$$

mit

$$p \in \{0,9999; 0,999; 0,99; 0,975; 0,95; 0,9; 0,75; 0,5; 0,25; 0,1; 0,05; 0,01\},$$

wobei die Berechnung der Verteilungsfunktion, wie durch die Notation $F_{Laguerre}^Q$ ausgedrückt werden soll, mit Hilfe von verallgemeinerten Laguerre-Polynomen erfolgte. Da es bei der Konfidenzintervallkonstruktion besonders wichtig ist, die Verteilungsfunktion von Q im Bereich großer Werte des Arguments q möglichst genau bestimmen zu können, wurde die Auswahl der Werte von p „unsymmetrisch“ vorgenommen. In Tabelle 3.2 sind die sich ergebenden Differenzen

$$10^6 \left(F_{Jacobi}^Q(q) - F_{Laguerre}^Q(q) \right) = 10^6 \left(F_{Jacobi}^Q(q) - p \right),$$

in Tabelle 3.3 die Differenzen

$$10^6 \left(F_{Laplace}^Q(q) - F_{Laguerre}^Q(q) \right) = 10^6 \left(F_{Laplace}^Q(q) - p \right)$$

vertafelt.

Die Resultate zeigen, daß es in jeder Situation mindestens zwei Verfahren gibt, die eine gute Übereinstimmung aufweisen, im Falle $n = 10$ und besonders für $n = 20$ gilt dies — mit Ausnahme sehr kleiner Werte von a — sogar für alle drei Verfahren, die sich somit sozusagen gegenseitig „verifizieren“. Deutlich zu sehen ist, daß die Verteilungsfunktionsbestimmung mittels Jacobi-Polynomen nicht bei extrem kleinen Werten des Parameters a angewendet werden darf, die Rücktransformation der Laplace-Transformierten kann hingegen nur für kleine Stichprobenumfänge verwendet werden, die Bestimmung der Verteilungsfunktion über verallgemeinerte Laguerre-Polynome liefert in allen Fällen gute Resultate.

$n=10: 10^6(F_{Jacobi}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	66	-346	3716	6319	-14325	-20980	-28647	184873	355392	535389	623229	734102
0,25	2	3	10	-49	49	195	287	542	1484	-2929	-1287	829
0,50	1	5	-7	5	-4	-10	9	5	-5	5	-3	2
1,00	0	1	-3	2	2	-3	4	-2	-2	0	-1	-1
2,50	0	0	-3	1	2	4	-3	2	-2	1	0	0
5,00	0	1	-1	-2	-4	0	4	-3	-5	2	16	-1
7,50	0	1	1	-1	-2	-4	2	4	-2	1	0	-1
10,00	0	0	1	0	-1	-2	0	2	-1	-2	0	0
25,00	0	0	1	1	1	0	-1	-3	0	0	-1	-1
100,00	0	0	0	0	0	0	0	0	-1	0	0	0

$n=20: 10^6(F_{Jacobi}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	-165	90	2293	-5319	-18315	-25084	12191	138497	281744	408671	489555	621465
0,25	0	1	1	0	-2	3	-4	-7	-1	4	-8	-2
0,50	0	0	0	0	0	0	0	0	0	0	0	0
1,00	0	0	0	0	0	0	0	0	0	0	0	0
2,50	0	0	0	0	1	2	-1	1	0	-1	0	0
5,00	0	0	0	-1	0	0	1	-1	0	0	0	0
7,50	0	0	0	0	-1	-1	1	1	-1	0	0	0
10,00	0	0	0	0	-1	-1	0	1	0	-1	-1	0
25,00	0	0	0	0	0	0	-1	0	0	0	1	0
100,00	0	0	0	0	0	0	0	0	0	0	0	0

$n=50: 10^6(F_{Jacobi}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	-6	-868	5620	4259	-2481	-26425	-34585	23568	118880	163631	169605	204744
0,25	0	0	-1	1	1	-3	1	15	-16	19	4	-6
0,50	0	0	0	0	0	0	0	0	0	0	0	0
1,00	0	0	0	0	0	0	0	0	0	0	0	0
2,50	0	0	0	0	0	0	0	0	0	0	0	0
5,00	0	0	0	0	-1	-2	0	1	0	-1	0	0
7,50	0	0	0	0	0	-1	0	0	0	0	0	0
10,00	0	0	0	0	0	-1	0	0	0	0	0	0
25,00	0	0	0	0	0	0	0	0	0	0	0	0
100,00	0	0	0	0	0	0	0	0	0	0	0	0

$n=250: 10^6(F_{Jacobi}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	1	206	-265	2903	-1438	8426	6154	-21806	-7798	-1116	7674	5294
0,25	0	0	0	0	0	0	0	0	0	1	0	0
0,50	0	0	0	0	0	0	0	0	0	0	0	0
1,00	0	0	0	0	0	0	0	0	0	0	0	0
2,50	0	0	0	0	0	0	0	0	0	0	0	0
5,00	0	0	0	0	0	0	0	0	0	0	0	0
7,50	0	0	0	0	0	0	0	0	0	1	0	0
10,00	0	0	0	0	0	0	-1	0	0	0	0	0
25,00	0	0	0	0	0	0	0	-1	-1	0	0	0
100,00	0	0	0	0	0	0	0	0	0	0	0	0

$n=1000: 10^6(F_{Jacobi}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	0	-3	51	23	-1	33	45	12	-38	-198	-32	210
0,25	0	0	0	0	1	0	-1	0	1	-1	0	0
0,50	0	0	0	0	0	0	0	0	0	-1	0	0
1,00	0	0	0	0	-1	0	1	1	0	0	0	0
2,50	0	0	0	0	0	0	0	0	0	0	0	0
5,00	0	0	0	0	0	0	0	0	0	0	0	0
7,50	0	0	0	0	0	0	-2	0	0	1	0	0
10,00	0	0	1	1	0	1	-1	-3	-2	1	1	1
25,00	0	0	0	0	0	0	0	0	0	0	0	0
100,00	0	0	0	0	0	0	0	0	0	0	0	0

Tab. 3.2 Zum numerischen Vergleich von F_{Jacobi}^Q mit $F_{Laguerre}^Q$

$n=10: 10^6(F_{Laplace}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	4	-6	15	-16	-28	26	-31	31	18	-14	8	-5
0,25	2	-4	8	18	-15	16	-18	18	-16	-9	5	-3
0,50	1	4	-8	5	-5	-9	9	6	-6	3	-3	-1
1,00	0	1	-2	1	3	-2	2	-3	-1	1	-1	-1
2,50	0	0	0	0	0	0	0	-1	-1	0	0	0
5,00	0	0	0	0	0	0	0	0	0	0	0	0
7,50	0	0	0	0	0	0	0	0	0	0	0	0
10,00	0	0	0	0	0	0	0	0	0	0	0	0
25,00	0	0	0	0	0	0	0	0	0	0	0	0
100,00	0	0	0	0	0	0	0	0	0	0	0	0

$n=20: 10^6(F_{Laplace}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	0	-1	0	1	-1	-2	2	-3	0	-2	0	0
0,25	0	0	-1	-1	1	2	1	-1	1	0	0	0
0,50	0	0	0	0	0	0	0	0	0	0	0	0
1,00	0	0	0	0	0	0	0	0	0	0	0	0
2,50	0	0	0	0	0	0	0	0	0	0	0	0
5,00	0	0	0	0	0	0	0	0	0	0	0	0
7,50	0	0	0	0	0	0	0	0	0	0	0	0
10,00	0	0	0	0	0	0	0	0	0	0	0	0
25,00	0	0	0	0	0	0	0	0	0	0	0	0
100,00	0	0	0	0	-1	0	0	0	0	1	0	0

$n=50: 10^6(F_{Laplace}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	-17	242	-622	-141	39	172	-369	-177	228	48	-405	-28
0,25	37	100	20	-22	-5	156	-21	-94	63	0	-36	1
0,50	-38	39	41	37	-44	1	14	-30	24	9	-19	-7
1,00	0	13	38	-4	-33	-2	39	-8	9	1	-83	5
2,50	1	9	4	-2	-11	4	18	-18	17	6	-10	25
5,00	3	10	30	0	-21	3	-2	-24	8	2	-13	3
7,50	3	7	29	-1	-16	6	22	-22	12	-30	-75	4
10,00	2	-11	7	-70	-8	5	13	-20	15	-30	-76	-3
25,00	3	4	5	2	-10	1	14	-19	14	-29	-78	-3
100,00	1	3	4	4	-10	3	19	-22	15	-7	-79	-29

$n=250: 10^6(F_{Laplace}^Q(q)-p)$												
a	p											
	0,9999	0,999	0,990	0,9750	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,010
0,10	-4235	28682	20633	9154	-4976	-16216	-153442	-1698	22322	14729	8086	-15590
0,25	-11779	38077	20947	11504	-4371	-10159	-146088	-1267	15717	12148	2488	-13596
0,50	-13614	13684	17785	12016	2508	-8047	-17066	-1233	13070	7645	315	-12153
1,00	-9731	10958	19260	11272	423	-9439	-22729	-1068	10946	4986	-183	-11540
2,50	-13375	39976	22370	12186	4603	-8081	-14294	-1178	18278	5365	8188	-12093
5,00	-8651	2424	12466	9200	4259	-7703	-20858	-935	18293	14371	8737	-11630
7,50	-4822	4350	20162	9414	4933	-7639	-13563	-927	18332	6058	8958	-11836
10,00	-10893	2251	20466	9521	5184	-3557	-13458	-1133	18356	14556	9083	-11956
25,00	-11676	10159	21024	9714	4446	-7863	-13516	-1179	18410	5294	9330	-12228
100,00	-5878	19860	12700	9825	5059	-4594	-20609	-918	18436	5767	9465	-9846

Tab. 3.3 Zum numerischen Vergleich von $F_{Laplace}^Q$ mit $F_{Laguerre}^Q$

4. Ein einseitiges Konfidenzintervall für die Varianz

4.1. Modellbeschreibung und Zielvorstellung

Zur Notation

In der Einleitung wurde die Fragestellung der Bestimmung einer oberen Schranke der Varianz beschrieben. Diese Abschätzung soll aufgrund einer Beobachtung $\underline{x} \stackrel{\text{def}}{=} (x_1, \dots, x_n)$ einer Stichprobe $\underline{X} \stackrel{\text{def}}{=} (X_1, \dots, X_n)$ vom Umfang n geschehen, wobei entsprechend des gewählten Modells die Annahme

$$(X_1, \dots, X_n) \text{ unabhängig, identisch } G(a, b)\text{-verteilt} \quad (4.1)$$

getroffen wird. Die Schranke soll im folgenden mit $c(\underline{X})$ bezeichnet werden. Gesucht ist ein einseitiges Konfidenzintervall, welches aus der Menge der σ^2 besteht, die kleiner oder gleich $c(\underline{X})$ sind, also

$$C(\underline{X}) \stackrel{\text{def}}{=} \{\sigma^2 : \sigma^2 \leq c(\underline{X})\}. \quad (4.2)$$

Realisationen der Zufallsvariablen $c(\underline{X})$ bzw. der zufälligen Menge $C(\underline{X})$ werden mit $c(\underline{x})$ bzw. $C(\underline{x})$ bezeichnet.

Das Konfidenzintervall $C(\underline{X})$ soll dabei so konstruiert werden, daß, gleichgültig welche Gammaverteilung vorliegt, also für beliebige positive Werte der Parameter a und b (bzw. a und σ^2), die Wahrscheinlichkeit, daß der wahre Wert der Varianz σ^2 vom Konfidenzintervall $C(\underline{X})$ überdeckt wird, eine vorgegebene Mindestwahrscheinlichkeit — das Konfidenzniveau — nicht unterschreitet. Das Konfidenzniveau wird mit $1 - \alpha$ ($0 < 1 - \alpha < 1$) bezeichnet. Die Niveaubedingung lautet also

$$\forall a > 0 \forall \sigma^2 > 0 : P_{a, \sigma^2} \{\sigma^2 \in C(\underline{X})\} \geq 1 - \alpha \quad (4.3)$$

wobei die Indizierung mit (a, σ^2) die „wahren“ Parameterwerte bezeichnen soll. Da aufgrund der Einseitigkeit (4.2) von $C(\underline{X})$, mit jedem Wert $\sigma_2^2 \in C(\underline{x})$ auch alle Werte σ_1^2 mit $\sigma_1^2 < \sigma_2^2$ in $C(\underline{x})$ liegen, gilt also

$$\begin{aligned} \forall a > 0 \forall \sigma_1^2 > 0 \forall \sigma_2^2 > 0 \text{ mit } \sigma_1^2 < \sigma_2^2 : \\ P_{a, \sigma_1^2} \{\sigma_1^2 \in C(\underline{X})\} \geq P_{a, \sigma_2^2} \{\sigma_2^2 \in C(\underline{X})\}. \end{aligned} \quad (4.4)$$

Zielvorstellung

Das Idealziel wäre die Konstruktion von $C(\underline{X})$ derart, daß in (4.3) stets das Gleichheitszeichen steht, also

$$\forall a > 0 \forall \sigma^2 > 0: P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \} = 1 - \alpha \quad (4.5)$$

gilt. Die Frage, ob es überhaupt ein $C(\underline{X})$ gibt, welches diese Idealvorstellung — die Ähnlichkeit — erfüllt, ist allerdings noch offen. Zudem wäre zu klären, ob sich ein solches Konfidenzintervall in der unter (4.2) gegebenen Form darstellen ließe oder ob sich eine für praktische oder rechnerische Belange ungeeignete Form ergäbe. Selbst nach Klärung dieser Fragen wäre es schon aus numerischen Gründen wohl dennoch unmöglich $C(\underline{X})$ so genau zu berechnen, daß die Beziehung (4.5) exakt erfüllt wäre.

Für praktische Belange ist es im allgemeinen ausreichend, für ein kleines, gegebenes $\epsilon > 0$ ein Konfidenzintervall $C(\underline{X})$ so angeben zu können, daß die Forderung

$$\forall a > 0 \forall \sigma^2 > 0: P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \} \in (1 - \alpha - \epsilon, 1 - \alpha + \epsilon) \quad (4.6)$$

erfüllt ist.

Aus der Diskussion der Gestalt der Gammaverteilungen (s. Kap. 2.1) ergibt sich zudem die Möglichkeit einer weiteren, aus praktischer Sicht unbedeutenden Abschwächung von (4.6). Verteilungen, deren Gestalt durch extrem kleine Werte von a gegeben sind, dürften wohl selten vorliegen, erst für $a > 1$ gilt für die Gammadichte $f_{a,b}(0) = 0$, für $a < 1$ besitzt die Dichtefunktion an der Stelle 0 einen Pol. Im weiteren werden Überdeckungswahrscheinlichkeiten stets für den Bereich $a \geq 0,1$ berechnet, Verteilungen, die eine noch extremere Gestalt besitzen, als die ohnehin schon extreme $G(0,1;b)$ -Verteilung werden also nicht mehr betrachtet. Obwohl nach den obigen Erläuterungen wohl schon ein sehr großer „Sicherheitsbereich“ gegeben ist, könnte — bei Bedarf — die Grenze 0,1 noch deutlich nach unten verschoben werden. Schließlich gibt es Hinweise, die eine theoretische Behandlung des Grenzüberganges $a \rightarrow 0$ nicht unmöglich erscheinen lassen (s. entsprechende Aussagen in den Kapiteln 2.3–2.7 sowie die in Kapitel 3.1.2 vorgestellte Rekursionsbeziehung, die den Grenzübergang $a \rightarrow 0$ bei der Gammaverteilung auf den Grenzübergang $a \rightarrow 1$ (Exponentialverteilung) zurückführt). Diese Fragen sollen an dieser Stelle — da wohl ausschließlich von theoretischem Interesse — nicht behandelt werden.

Wesentlich gravierender ist die Einschränkung, daß das hier vorgestellte Verfahren zur Konstruktion von $C(\underline{X})$ die Überdeckungswahrscheinlichkeiten lediglich für endlich viele Werte des Parameters a kontrollieren kann. Zwar wurde eine große Menge A von Werten des Parameters a aus einem großen Bereich (0,1–1000) gewählt, dennoch stellt die Bedingung (4.7)

Konstruiere $C(\underline{X})$, so daß für ein gegebenes $\epsilon > 0$ gilt:

$$\forall a \in A \forall \sigma^2 > 0: P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \} \in (1 - \alpha - \epsilon, 1 - \alpha + \epsilon) \quad (4.7)$$

mit der in Tabelle 4.1 angegebenen Stützstellenmenge

$$A = \{a_1, \dots, a_{97}\}$$

eine deutliche Einschränkung gegenüber der Forderung (4.6) dar.

$A = \{$	0,1 ;	0,125;	0,15 ;	0,175;	0,2 ;	0,25 ;	0,3 ;
	0,35 ;	0,4 ;	0,45 ;	0,5 ;	0,625;	0,75 ;	0,875;
	1 ;	1,125;	1,25 ;	1,375;	1,5 ;	1,625;	1,75 ;
	1,875;	2 ;	2,2 ;	2,4 ;	2,6 ;	2,8 ;	3 ;
	3,2 ;	3,4 ;	3,6 ;	3,8 ;	4 ;	4,2 ;	4,4 ;
	4,6 ;	4,8 ;	5 ;	5,5 ;	6 ;	6,5 ;	7 ;
	7,5 ;	8 ;	8,5 ;	9 ;	9,5 ;	10 ;	11 ;
	12 ;	13 ;	14 ;	15 ;	16 ;	17 ;	18 ;
	19 ;	20 ;	21 ;	22 ;	23 ;	24 ;	25 ;
	27,5 ;	30 ;	32,5 ;	35 ;	37,5 ;	40 ;	42,5 ;
	45 ;	47,5 ;	50 ;	55 ;	60 ;	65 ;	70 ;
	75 ;	80 ;	90 ;	100 ;	125 ;	150 ;	175 ;
	200 ;	225 ;	250 ;	300 ;	350 ;	400 ;	450 ;
	500 ;	600 ;	700 ;	800 ;	900 ;	1000 ;	}

Tab. 4.1 Stützstellenmenge A

Um diese Schwäche des Verfahrens auszugleichen, wird in Kapitel 5 ein Verfahren angegeben, welches es ermöglicht, das aufgrund der großen Zahl von Stützstellen zu erwartende „Wohlverhalten“ der Überdeckungswahrscheinlichkeiten für $a \notin A$ zu überprüfen, indem Abschätzungen der Überdeckungswahrscheinlichkeiten berechnet werden. Diese Abschätzungen können allerdings ein gegebenes Konfidenzintervall nur nachträglich kontrollieren, während die Stützstellenmenge A über die Forderung (4.7) zur Konstruktion von $C(\underline{X})$ herangezogen wird.

Schon aus numerischen Gründen, aber auch aus Gründen, die in der Konstruktionsvorschrift des Konfidenzintervalles $C(\underline{X})$ liegen, ist es nicht möglich, ein beliebig kleines ϵ vorzugeben. In allen hier berechneten Fällen (Stichprobenumfänge zwischen 10 und 250 sowie Konfidenzniveaus 0,90; 0,95 und 0,99) war es jedoch möglich, beliebige Werte $\epsilon \geq 0,00025$ vorzugeben. Auch dieser Bereich läßt sich — bei entsprechend größerem Rechenaufwand — noch deutlich vergrößern.

4.2. Die Form des Konfidenzintervalls

Mit Hilfe von zwei — wohl unproblematischen — einschränkenden Annahmen kann die Funktion c genauer spezifiziert werden.

1. Annahme:

c soll eine Funktion des für die Gamma-Verteilung suffizienten Tupels $(\hat{a}_{ML}, \hat{\sigma}_{ML}^2)$ sein, also

$$c(\underline{X}) = c(\hat{a}_{ML}, \hat{\sigma}_{ML}^2). \quad (4.8)$$

Ausgeschlossen sind damit Ansätze, die etwa die Stichprobenvarianz $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ einbeziehen, sowie die meisten Ansätze, die andere — in gewissem Sinne robuste Varianzschätzer — verwenden. Die zweite Annahme soll über die Dualität Testfamilie–Konfidenzintervall, die vorbereitend beschrieben wird, eingeführt werden.

Dualität Testfamilie–Konfidenzintervall

Im Rahmen der statistischen Inferenz wird häufig ein Konfidenzintervall aus einer Klasse von Tests hergeleitet. Dieser „Umweg“, der ausführlich von Lehmann (1986, S. 213ff) dargestellt wird, erweist sich als günstig, da er es erlaubt, Begriffe und Ergebnisse der klassischen Neyman-Pearson Testtheorie auf Konfidenzintervalle zu übertragen.

Sei zu jedem positiven σ_0^2 zum einseitigen Testproblem

$$H_{\sigma_0^2}^0 : \sigma^2 \geq \sigma_0^2, a > 0, \quad H_{\sigma_0^2}^1 : \sigma^2 < \sigma_0^2, a > 0$$

je ein nichtrandomisierter Test

$$\Phi_{\sigma_0^2}(\underline{X})$$

zum Signifikanzniveau α gegeben:

$$\forall a > 0 \quad \forall \sigma^2 \geq \sigma_0^2 : P_{a, \sigma^2} \left\{ \Phi_{\sigma_0^2}(\underline{X}) = 1 \right\} \leq \alpha, \quad (4.9)$$

wobei die übliche Darstellung

$$\Phi_{\sigma_0^2}(\underline{x}) = \begin{cases} 1 & \text{Ablehnung von } H_{\sigma_0^2}^0 \\ 0 & \text{Nichtablehnung von } H_{\sigma_0^2}^0 \end{cases}$$

gelten soll. Die Familie der Tests $\left\{ \Phi_{\sigma_0^2}(\underline{X}) : \sigma_0^2 > 0 \right\}$ definiert ein Konfidenzintervall $\tilde{C}(\underline{X})$ für die Varianz σ^2 , indem alle Werte von σ_0^2 , für welche keine Ablehnung von $H_{\sigma_0^2}^0$ erfolgt, ins Konfidenzintervall $\tilde{C}(\underline{X})$ aufgenommen werden:

$$\tilde{C}(\underline{X}) = \left\{ \sigma_0^2 : \Phi_{\sigma_0^2}(\underline{X}) = 0 \right\} \quad (4.10)$$

Aus der Bedingung (4.9) folgt sofort

$$\forall a > 0 \quad \forall \sigma^2, \sigma_1^2 > 0 \text{ mit } \sigma_1^2 \leq \sigma^2 : P_{a, \sigma^2} \left\{ \sigma_1^2 \in \tilde{C}(\underline{X}) \right\} \geq 1 - \alpha. \quad (4.11)$$

Die Niveaubedingung (4.11) liegt — analog zur Niveaubedingung (4.9) für einen einseitigen Test — der Definition des einseitigen Konfidenzintervalls zugrunde, so daß das gesuchte Konfidenzintervall $C(\underline{X})$ aufgrund der Beziehungen (4.3) und (4.4) auch im Sinne dieser Definition einseitig ist. Ebenso wie ein einseitiges Testproblem keineswegs eine einseitige kritische Region ergeben muß, muß $\tilde{C}(\underline{X})$ nicht die hier geforderte Gestalt (4.2) besitzen, so daß eine zusätzliche Forderung an die Testfamilie $\{\Phi_{\sigma_0^2}(\underline{X}) : \sigma_0^2 > 0\}$ notwendig ist:

$$\forall \underline{x} \quad \forall \sigma_0^2 > 0 : \quad \begin{cases} \Phi_{\sigma_0^2}(\underline{x}) = 0 & \implies \forall \sigma_1^2 < \sigma_0^2 : \Phi_{\sigma_1^2}(\underline{x}) = 0 \\ \Phi_{\sigma_0^2}(\underline{x}) = 1 & \implies \forall \sigma_1^2 > \sigma_0^2 : \Phi_{\sigma_1^2}(\underline{x}) = 1 \end{cases} \quad (4.12)$$

Diese Bedingung ist einleuchtend, da aus der Ablehnung der Hypothese $H_{\sigma_0^2}^0$ auch die Ablehnung jeder Teilhypothese $H_{\sigma_1^2}^0$ ($\sigma_1^2 > \sigma_0^2$) folgen sollte, sowie aus der Nichtablehnung von $H_{\sigma_0^2}^0$ auch die Nichtablehnung jeder Oberhypothese $H_{\sigma_1^2}^0$ ($\sigma_1^2 < \sigma_0^2$). Damit bleibt noch die Freiheit, eine Familie von Tests $\{\Phi_{\sigma_0^2}(\underline{X}) : \sigma_0^2 > 0\}$ zu wählen, für welche (4.9) und (4.12) erfüllt sind. Die zweite Annahme wird diese „Freiheit“ auf die Wahl eines einzigen Tests zurückführen.

Für jedes $\sigma_0^2 > 0$ läßt sich die oben beschriebene Situation des Tests der Hypothese $H_{\sigma_0^2}^0$ gegen $H_{\sigma_0^2}^1$ durch eine einfache Transformation auf die Fragestellung des Tests von H_1^0 gegen H_1^1 zurückführen. Dazu wird die Stichprobe $(X_1, \dots, X_n)$ gemäß

$$Y_i \stackrel{\text{def}}{=} \frac{1}{\sigma_0} X_i, \quad i = 1, \dots, n. \quad (4.13)$$

transformiert. Die Zufallsvariablen $Y_1, \dots, Y_n$ sind unabhängig und identisch $G(a, \frac{\sigma_0^2}{\sigma_0^2})$ -verteilt, so daß unter $H_{\sigma_0^2}^0$ die Verteilung der Y_i eine Varianz besitzt, die größer oder gleich 1 ist, während unter $H_{\sigma_0^2}^1$ der Wert der Varianz kleiner 1 ist.

2. Annahme

Aus einem vorgegebenen Test $\Phi_1(\underline{X})$ für H_1^0 gegen H_1^1 wird die Familie $\{\Phi_{\sigma_0^2}(\underline{X}) : \sigma_0^2 > 0\}$ durch die Anwendung von Φ_1 auf $Y_1, \dots, Y_n$ definiert:

$$\Phi_{\sigma_0^2}(\underline{X}) \stackrel{\text{def}}{=} \Phi_1\left(\frac{1}{\sigma_0} \underline{X}\right) \quad (4.14)$$

Damit übertragen sich alle Eigenschaften des Tests $\Phi_1(\underline{X})$ direkt auf das daraus konstruierte Konfidenzintervall. Das unter (4.5) beschriebene Idealziel entspricht hier dem Ziel, $\Phi_1(\underline{X})$ unter Berücksichtigung von (4.12) so zu konstruieren, daß gilt

$$\forall a > 0 : \quad P_{a,1} \{\Phi_1(\underline{X}) = 0\} = \alpha.$$

Als letzte Beziehung wird noch die in Kapitel 2.3 hergeleitete Skalierungsinvarianz von Q und seiner Transformation $\hat{a}_{ML}$ benötigt:

Für alle $\underline{x}$ und alle positiven d gilt

$$\begin{aligned}\hat{a}(d\underline{x}) &= \hat{a}(\underline{x}) \\ \hat{\sigma}^2(d\underline{x}) &= d^2 \hat{\sigma}^2(\underline{x}).\end{aligned}\tag{4.15}$$

Aus den auf (4.2), (4.10) und (4.14) basierenden Umformungen

$$\begin{aligned}\sigma^2 \in C(\underline{x}) &\iff \sigma^2 \leq c(\underline{x}) \\ &\iff \Phi_{\sigma^2}(\underline{x}) = 0 \\ &\iff \Phi_1\left(\frac{1}{\sigma}\underline{x}\right) = 0 \\ &\iff \Phi_1\left(\frac{1}{d\sigma}d\underline{x}\right) = 0 \\ &\iff \Phi_{d^2\sigma^2}(d\underline{x}) = 0 \\ &\iff d^2\sigma^2 \leq c(d\underline{x}) \\ &\iff d^2\sigma^2 \in C(d\underline{x})\end{aligned}$$

ergibt sich aus der Äquivalenz $\sigma^2 \leq c(\underline{x}) \iff d^2\sigma^2 \leq c(d\underline{x})$

$$c(d\underline{x}) = d^2 c(\underline{x})$$

und mit (4.8) und (4.15) folgt

$$c(\hat{a}, d^2 \hat{\sigma}^2) = d^2 c(\hat{a}, \hat{\sigma}^2).$$

Die Wahl von $d = \frac{1}{\hat{\sigma}}$ und die Neudefinition von c

$$c(\hat{a}) \stackrel{\text{def}}{=} c(\hat{a}, 1)$$

liefert damit die endgültige Form des Konfidenzintervalls:

$$C(\underline{X}) = C(\hat{a}_{ML}, \hat{\sigma}_{ML}^2) = \left\{ \sigma^2 : \sigma^2 \leq c(\hat{a}_{ML}) \hat{\sigma}_{ML}^2 \right\}.\tag{4.16}$$

Zur Notation

Wie in Kapitel 2 bereits eingeführt, soll die Verteilungsfunktion der $G(a, \sigma^2)$ -Verteilung im folgenden stets durch das Symbol

$$\mathcal{G}_{a, \sigma^2},$$

die Verteilungsfunktion von $\hat{a}_{ML}$ durch

$$\mathcal{A}_{a, n}$$

bezeichnet werden. Aus der Beziehung

$$X \sim G(a, \sigma^2) \quad \Rightarrow \quad \frac{X}{\sigma} \sim G(a, 1)$$

ergibt sich

$$\mathcal{G}_{a, \sigma^2}(x) = \mathcal{G}_{a, 1}\left(\frac{x}{\sigma}\right).$$

Der Kürze halber soll $\mathcal{G}_{a, 1}$ stets mit $\mathcal{G}_a$ bezeichnet werden, so daß sich für die Verteilungsfunktion von $\bar{X}$ die Darstellung

$$P_{a, \sigma^2} \{ \bar{X} \leq x \} = \mathcal{G}_{a n, \frac{\sigma^2}{n}}(x) = \mathcal{G}_{a n} \left(x \sqrt{\frac{n}{\sigma^2}} \right) \quad (4.17)$$

ergibt.

Überdeckungswahrscheinlichkeit

Die Wahrscheinlichkeit, daß bei wahren (a, σ^2) das Konfidenzintervall $C(\underline{X})$ den Wert σ_1^2 überdeckt

$$\begin{aligned} P_{a, \sigma^2} \{ \sigma_1^2 \in C(\underline{X}) \} &= P_{a, \sigma^2} \left\{ \sigma_1^2 \leq c(\hat{a}_{ML}) \hat{\sigma}_{ML}^2 \right\} \\ &= P_{a, \sigma^2} \left\{ \sigma_1^2 \leq \frac{c(\hat{a}_{ML})}{\hat{a}_{ML}} X^2 \right\} \\ &= 1 - P_{a, \sigma^2} \left\{ \bar{X} \leq \sqrt{\sigma_1^2} \sqrt{\frac{\hat{a}_{ML}}{c(\hat{a}_{ML})}} \right\} \end{aligned}$$

läßt sich unter Verwendung des Stieltjes-Integralbegriffs und unter Berücksichtigung der Unabhängigkeit von $\bar{X}$ und $\hat{a}_{ML}$ folgendermaßen darstellen:

$$\begin{aligned} P_{a, \sigma^2} \{ \sigma_1^2 \in C(\underline{X}) \} &= 1 - \int_0^\infty \mathcal{G}_{a n, \frac{\sigma^2}{n}} \left(\sqrt{\sigma_1^2} \sqrt{\frac{\hat{a}}{c(\hat{a})}} \right) d\mathcal{A}_{a, n}(\hat{a}) \\ &= 1 - \int_0^\infty \mathcal{G}_{a n} \left(\sqrt{\frac{\sigma_1^2}{\sigma^2}} \sqrt{\frac{\hat{a} n}{c(\hat{a})}} \right) d\mathcal{A}_{a, n}(\hat{a}). \end{aligned} \quad (4.18)$$

Es zeigt sich also, daß die Überdeckungswahrscheinlichkeit nur von a sowie vom Quotienten $\frac{\sigma^2}{a^2}$ abhängt, so daß durch die für positive a und ν definierte Funktion

$$\tilde{\alpha}(a, \nu) \stackrel{\text{def}}{=} \int_0^\infty \mathcal{G}_{an} \left(\sqrt{\nu \frac{\hat{a}n}{c(\hat{a})}} \right) d\mathcal{A}_{a,n}(\hat{a}) \quad (4.19)$$

die Kennfunktion

$$P_{a,\sigma^2} \{ \nu \sigma^2 \in C(\underline{X}) \} = 1 - \tilde{\alpha}(a, \nu)$$

bestimmt ist. Als Funktion von ν (bei festem a) aufgefaßt, ist $\tilde{\alpha}(a, \cdot)$ antiton (s. (4.4)). Da im folgenden $\tilde{\alpha}$ stets als Funktion von a bei festem Wert $\nu = 1$ verwendet wird, soll die folgende Kurzschreibweise

$$\tilde{\alpha}(\cdot) \stackrel{\text{def}}{=} \tilde{\alpha}(\cdot, 1)$$

eingeführt werden.

Die Zielvorstellung (4.7) lautet damit:

Bestimme $C(\underline{X})$, so daß für ein vorgegebenes positives ϵ gilt

$$\forall a \in A : \quad \alpha - \epsilon < \tilde{\alpha}(a) < \alpha + \epsilon. \quad (4.20)$$

Die Familie $\{h_a : a > 0\}$

Aus der Formel (4.19) ergibt sich

$$\tilde{\alpha}(a) = \int_0^\infty \mathcal{G}_{an} \left(\sqrt{\frac{\hat{a}n}{c(\hat{a})}} \right) d\mathcal{A}_{a,n}(\hat{a})$$

und mit der Substitution $u \stackrel{\text{def}}{=} \mathcal{A}_{a,n} \circ \hat{a}$

$$\tilde{\alpha}(a) = \int_0^1 \mathcal{G}_{an} \left(\sqrt{\frac{n\mathcal{A}_{a,n}^{-1}(u)}{c(\mathcal{A}_{a,n}^{-1}(u))}} \right) du,$$

wobei $\mathcal{A}_{a,n}^{-1}$ die Umkehrfunktion von $\mathcal{A}_{a,n}$ bezeichnet.

Mit der für $u \in (0, 1)$ definierten Funktion

$$h_a : \quad u \quad \longmapsto \quad \mathcal{G}_{an} \left(\sqrt{\frac{n\mathcal{A}_{a,n}^{-1}(u)}{c(\mathcal{A}_{a,n}^{-1}(u))}} \right) \quad (4.21)$$

ergibt sich

$$\tilde{\alpha}(a) = \int_0^1 h_a(u) du. \quad (4.22)$$

Die zweistufige Abbildung

$$h_a \circ u: \hat{a} \mapsto u = \mathcal{A}_{a,n}(\hat{a}) \mapsto h_a(u) = \mathcal{G}_{an} \left(\sqrt{\frac{n\mathcal{A}_{a,n}^{-1}(u)}{c(\mathcal{A}_{a,n}^{-1}(u))}} \right) \quad (4.23)$$

verdeutlicht die wichtige Tatsache, daß bei der Integration von h_a bei festem a im Endeffekt über $\hat{a}$ integriert wird. Aus dieser Abbildung läßt sich auch der folgende Zusammenhang erkennen:

$$\forall \hat{a} > 0: P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \mid \hat{a}_{ML} = \hat{a} \} = 1 - h_a \circ \mathcal{A}_{a,n}(\hat{a}) \quad (4.24)$$

Monotonie

Aus der Monotonie der Inversen $\mathcal{A}_{a,n}^{-1}$ ergeben sich die folgenden Eigenschaften der Funktion h_a :

Ist die positive Funktion c im Bereich $\mathbb{R}^+$ stetig und antiton, so gilt: Für alle $a > 0$ ist die Funktion h_a eine bijektive, streng isotone Abbildung aus der Menge $(0, 1)$ in die Menge $(0, 1)$.

Die in diesem Kapitel vorgeschlagenen Konfidenzintervalle, die die Zielvorgabe (4.7) erfüllen, werden typischerweise durch Funktionen c , die diese Eigenschaften besitzen, gebildet. Damit ergibt sich für solche Konfidenzintervalle, daß die bedingte Überdeckungswahrscheinlichkeit $P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \mid \hat{a}_{ML} = \hat{a} \}$, als Funktion von $\hat{a}$ aufgefaßt, streng antiton ist mit $\lim_{\hat{a} \rightarrow 0} P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \mid \hat{a}_{ML} = \hat{a} \} = 1$ und $\lim_{\hat{a} \rightarrow \infty} P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \mid \hat{a}_{ML} = \hat{a} \} = 0$. In Abschnitt 4.6 wird gezeigt, was passiert, wenn versucht wird ein Konfidenzintervall zu konstruieren, für welches die bedingten Überdeckungswahrscheinlichkeiten konstant α sind.

4.3. Das Konstruktionsprinzip

Zur Problematik der numerischen Berechnung von c

Die Berechnung der Funktion c kann in einem iterativen Verfahren, dessen Grundzüge hier beschrieben werden sollen, durchgeführt werden. Bei dieser Konstruktion ergeben sich drei Probleme:

- I Auf welche Weise kann eine geeignete Startfunktion $c^{(1)}$ vorgegeben werden?
- II Wie kann aus der im i -ten Iterationsschritt bestimmten Funktion $c^{(i)}$ die Funktion $\hat{c}^{(i)}$ berechnet werden?
- III Wie kann $c^{(i+1)}$ aus $c^{(i)}$ und $\hat{c}^{(i)}$ konstruiert werden?

Die Fragen I und III werden im folgenden im Vordergrund stehen. Die Frage III ist dabei von besonderer Bedeutung, da dieser „Korrekturmechanismus“ schließlich die Erfüllung der Zielvorgabe (4.7) gewährleisten soll. Das Problem II ist ein rein numerisches Problem und wird an geeigneter Stelle behandelt. Im wesentlichen

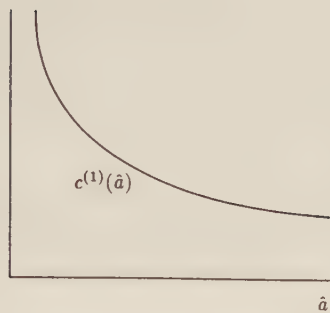


Abb. 4.1 Die Funktion $c^{(1)}$

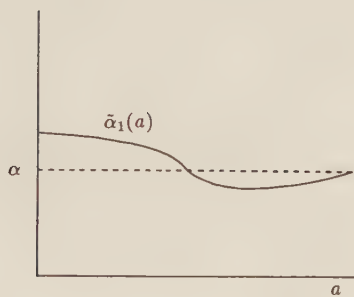


Abb. 4.2 Die Funktion $\tilde{\alpha}^{(1)}$

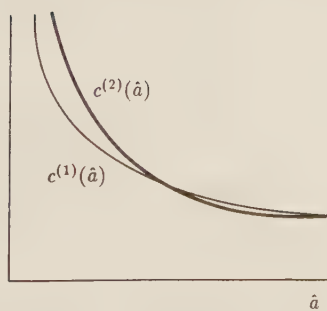


Abb. 4.3 Die Funktionen $c^{(1)}$ und $c^{(2)}$

handelt es sich um die Frage der Berechnung der Verteilungsfunktionen von $\bar{X}$ und Q (s. Kapitel 3) und der numerischen Bestimmung des Integrals (4.20). Stets wird die Funktion $c^{(i)}$ nur an einer endlichen Menge von Stützstellen $\hat{a} \in \hat{A}$ berechnet und mittels eines geeigneten Interpolationsalgorithmus, der in Abschnitt 4.9 vorgestellt wird, auf den Bereich $\hat{a} \notin \hat{A}$ übertragen. Die in den Beispielen dieser Arbeit stets verwendete Stützstellenmenge $\hat{A}$ besteht aus sämtlichen Elementen der Menge A (s. Tabelle 4.1) sowie den sieben zusätzlichen Werten $\{0,05; 0,0625; 0,075; 1250; 1500; 1750; 2000\}$.

Anhand eines hypothetischen Iterationsschritts von $c^{(1)}$ nach $c^{(2)}$, der in den Abbildungen 4.1–4.3 illustriert wird, soll eine Grundidee der in Abschnitt 4.9 ausführlich behandelten Lösung des Problems III vermittelt werden (aus zeichnerischen Gründen werden die Funktionen auch außerhalb ihrer Stützstellenbereiche angegeben).

Im Bereich kleiner Werte von a liegt die Funktion $\tilde{\alpha}^{(1)}$ oberhalb der Konstanten α , die Überdeckungswahrscheinlichkeit $1 - \tilde{\alpha}^{(1)}$ damit unterhalb des Konfidenzniveaus $1 - \alpha$, das Konfidenzintervall ist also zu „kurz“. Im Bereich mittlerer Werte von a trifft die umgekehrte Aussage zu, das Konfidenzintervall ist zu „lang“, während für sehr große Werte des Parameters a die Überdeckungswahrscheinlichkeit den Wert $1 - \alpha$ näherungsweise erreicht. Da $\hat{a}_{ML}$ ein Schätzer für a ist, wird eine Korrektur, die eine bessere Einhaltung des Konfidenzniveaus zum Ziel hat, die in Abb. 4.3 angedeutete Form besitzen. Im Bereich kleiner Werte von $\hat{a}$ wird $c^{(2)}$ größer als $c^{(1)}$ sein, für mittlere $\hat{a}$ kleiner, und für große Werte von $\hat{a}$ werden sich $c^{(2)}$ und $c^{(1)}$ nur noch geringfügig unterscheiden. Wird das Ausmaß der Korrektur, wie in Abschnitt 4.9 erläutert, geeignet gewählt, so wird $\tilde{\alpha}^{(2)}$ der Zielvorstellung (4.7) besser entsprechen als $\tilde{\alpha}^{(1)}$ (die in der Abb. 4.3 skizzierte Korrektur ist sicherlich „ungeeignet“, in Wirklichkeit sind — wie in Abschnitt 4.9 demonstriert — nur winzige Korrekturen nötig).

Zur numerischen Berechnung von $\tilde{\alpha}$

Im folgenden sollen einige Überlegungen zur numerischen Bestimmung des Integrals (4.22) angestellt werden. Die numerische Berechnung der Verteilungsfunktion $\mathcal{A}_{a,n}$ wird ausführlich in Kapitel 3 behandelt. Aus der isotonen Funktion $\mathcal{A}_{a,n}$ wird mit Standardverfahren der numerischen Mathematik die Inverse $\mathcal{A}_{a,n}^{-1}$ bestimmt, so daß sich für jedes $u \in (0, 1)$ der Funktionswert $h_a(u)$ gemäß der Formel (4.21) berechnen läßt.

Durch die geeignete Wahl einer Stützstellenmenge $\{u_1, \dots, u_k\}$ kann $\tilde{\alpha}(a)$ sodann mit hoher Genauigkeit bestimmt werden. Am Beispiel einiger Funktionen aus einer solchen Familie $\{h_a : a > 0\}$ lassen sich leicht Regeln für eine sinnvolle Stützstellenwahl ableiten. In Abbildung 4.4 sind drei Funktionen h_{a_1}, h_{a_2} und h_{a_3} als typische Repräsentanten für kleine, mittlere und große Werte von a gezeichnet (diese Funktionen ergeben sich aus der Startfunktion $c^{(1)}$ des Beispiels $n = 50$, $1 - \alpha = 0,95$ für $(a_1, a_2, a_3) = (0,5; 5; 100)$).

Es fällt auf, daß alle drei Funktionen im Bereich kleinerer sowie mittlerer Werte von u kaum ansteigen (lediglich für $a \ll a_1$ steigt h_a im Bereich sehr kleiner

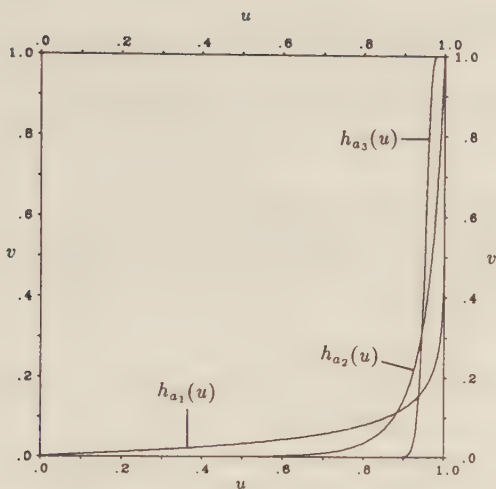


Abb. 4.4 Die Funktionen h_{a_1} , h_{a_2} und h_{a_3}

Werte von u geringfügig). Erst im Bereich großer Werte von u ist für alle drei Funktionen ein merklicher Anstieg zu verzeichnen, wobei die Funktion h_{a_3} in einem sehr kleinen Intervall um $u = 1 - \alpha$ besonders steil ansteigt. Um in allen drei Fällen den bei der numerischen Integration auftretenden Fehler gering zu halten, empfiehlt es sich, die Abstände der Stützstellen im Bereich großer und sehr kleiner Werte von u gering zu wählen, während in den Bereichen geringen Anstiegs auch relativ große Abstände keinen nennenswerten Fehler bewirken. Die den Beispielen dieser Arbeit zugrundeliegende Stützstellenmenge (238 Stützstellen) ist in Abb. 4.5 dargestellt. Ein Beispiel für die numerischen Auswirkungen dieser Wahl wird in Kapitel 5 gegeben.



Abb. 4.5 Die Stützstellenmenge $\{u_1, \dots, u_{238}\}$

4.4. Die indirekte Bestimmung von $C(\underline{X})$

Der Weg der direkten Bestimmung von $C(\underline{X})$ ist schwierig. Einfacher ist es, in einem ersten Schritt ein gemeinsames Konfidenzintervall für beide Parameter zu konstruieren und dieses dann auf $C(\underline{X})$ zu projizieren. Das gemeinsame

Konfidenzintervall $C^*(\underline{X})$ soll die Form

$$C^*(\underline{X}) = \{(a, \sigma^2) : \sigma^2 \leq c^*(a, \underline{X})\} \quad (4.25)$$

besitzen, also einseitig in σ^2 sein und die Niveaubedingung

$$\forall a > 0 \forall \sigma^2 > 0 : P_{a, \sigma^2} \{(a, \sigma^2) \in C^*(\underline{X})\} \geq 1 - \alpha \quad (4.26)$$

erfüllen.

Dualität Testfamilie-Konfidenzintervall

Ein solches Konfidenzintervall ist wiederum aus der Dualität Testfamilie-Konfidenzintervall konstruierbar. Aus einer Testfamilie $\{\Phi_{a_0, \sigma_0^2}(\underline{X}) : a_0 > 0, \sigma_0^2 > 0\}$ zur Familie der Hypothesen

$$H_{a_0, \sigma_0^2}^0 : a = a_0, \sigma^2 \geq \sigma_0^2 \quad H_{a_0, \sigma_0^2}^1 : a > 0, \sigma^2 < \sigma_0^2, \quad a_0 > 0, \sigma_0^2 > 0$$

wird in der üblichen Weise

$$C^*(\underline{x}) \stackrel{\text{def}}{=} \{(a, \sigma^2) : \Phi_{a, \sigma^2}(\underline{x}) = 0\}$$

gebildet. Diese Hypothesen decken zwar nicht den gesamten Raum der Gamma-Verteilungen ab, doch dies ist ohne Bedeutung (die nicht von $H_{a_0, \sigma_0^2}^0$ und $H_{a_0, \sigma_0^2}^1$ abgedeckten Verteilungen könnten ohne weiteres zu $H_{a_0, \sigma_0^2}^1$ hinzugefügt werden). Die in Abschnitt 4.2 erläuterten Annahmen können in völlig analoger Form auf diese Situation übertragen werden, woraus sich die Darstellung

$$C^*(\underline{X}) = \{(a, \sigma^2) : \sigma^2 \leq c^*(a, \hat{a}_{ML}) \hat{\sigma}_{ML}^2\}$$

ergibt. Die konkrete Bestimmung der zur Konstruktion von $C^*(\underline{X})$ benötigten Testfamilie $\{\Phi_{a, 1} : a > 0\}$ (in Analogie zum Test $\Phi_1(\underline{X})$ aus Abschnitt 4.2) läßt sich allerdings wiederum nur für eine endliche Zahl von Werten von a durchführen.

Projektion

$C^*(\underline{X})$ wird mit Hilfe der Projektionsvorschrift

$$\forall \sigma^2 > 0 : \sigma^2 \in C(\underline{X}) \iff \exists a > 0 : (a, \sigma^2) \in C^*(\underline{X}) \quad (4.27)$$

auf $C(\underline{X})$ abgebildet. Wenn es zu einem Wert σ^2 also auch nur ein einziges Paar (a, σ^2) in $C^*(\underline{X})$ gibt, so wird σ^2 in $C(\underline{X})$ aufgenommen. Diese Projektionsvorschrift kann durch die Abbildung 4.6 verdeutlicht werden. Für eine Realisation $(\hat{a}, \hat{\sigma}^2)$ der Zufallsvariablen $(\hat{a}_{ML}, \hat{\sigma}_{ML}^2)$ ergibt sich $c(\hat{a})$ aus $c^*(a, \hat{a})$ durch Projektion auf die Ordinatenachse.

Die Projektionsvorschrift (4.27) ist also gleichbedeutend mit der Definition

$$c(\hat{a}) \stackrel{\text{def}}{=} \sup_{a > 0} c^*(a, \hat{a}) \quad (4.28)$$

und führt unmittelbar zur Ungleichung

$$\forall a > 0 \forall \sigma^2 > 0 : P_{a, \sigma^2} \{\sigma^2 \in C(\underline{X})\} \geq P_{a, \sigma^2} \{(a, \sigma^2) \in C^*(\underline{X})\} \quad (4.29)$$

und damit aufgrund von (4.26) zur Erfüllung der Niveaubedingung durch $C(\underline{X})$.

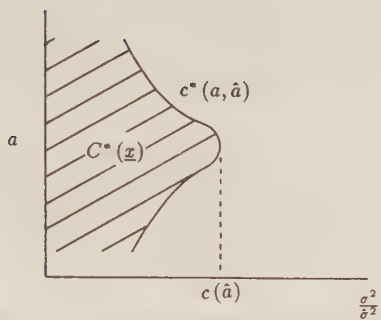


Abb. 4.6 Die Projektion von $C^*(\underline{x})$ auf $C(\underline{x})$

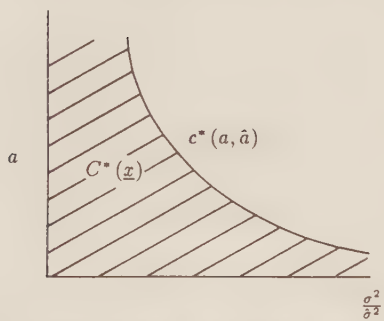


Abb. 4.7 Maximaler Projektionsverlust

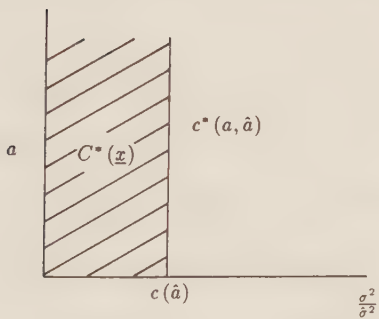


Abb. 4.8 Kein Projektionsverlust

Projektionsverlust

In der obigen Ungleichung (4.29) wird im allgemeinen ein „>“-Zeichen stehen. Es wird sich zeigen, daß es relativ leicht ist, $C^*(\underline{X})$ so zu konstruieren, daß in (4.26) ein „=“-Zeichen steht, die Konstruktion von $C^*(\underline{X})$, so daß auch in (4.29) das „=“-Zeichen steht, ist wesentlich schwieriger.

Falls für eine Realisation $(\hat{a}, \hat{\sigma}^2)$ dagegen $c^*(a, \hat{a})$ die in Abbildung 4.7 gezeichnete Gestalt besitzt, so lautet die Realisation von $C(\underline{X})$ offensichtlich $C(\underline{x}) = \{\sigma^2 : \sigma^2 < \infty\}$. Besitzt $C^*(\underline{X})$ mit Wahrscheinlichkeit 1 eine solche Gestalt (in den nächsten beiden Abschnitten werden zwei solche — keineswegs abwegige — Konstruktionen angegeben werden), so ist der „Projektionsverlust“ maximal, gleichgültig welche Optimalitätskriterien $C^*(\underline{X})$ auch immer erfüllen mag.

Besitzt $C^*(\underline{X})$ dagegen für alle Realisationen von $\hat{a}_{ML}$ die in Abbildung 4.8 angegebene Gestalt, so fällt keinerlei Projektionsverlust an, denn es gilt

$$\forall a > 0 \forall \sigma^2 > 0 : (a, \sigma^2) \in C^*(\underline{X}) \iff \sigma^2 \in C(\underline{X})$$

und damit

$$\forall a > 0 \forall \sigma^2 > 0 : P_{a, \sigma^2} \{ (a, \sigma^2) \in C^*(\underline{X}) \} = P_{a, \sigma^2} \{ \sigma^2 \in C(\underline{X}) \}. \quad (4.30)$$

Das in den nächsten Abschnitten verfolgte Ziel wird es daher sein, $C^*(\underline{X})$ so zu konstruieren, daß für alle $\hat{a} \in \hat{A}$ $c^*(a, \hat{a})$ die in Abbildung 4.8 gezeichnete Form besitzt. Diese Form soll durch die Sprechweise

„kein Projektionsverlust bei Beobachtung $\hat{a}$ “

zum Ausdruck gebracht werden.

4.5. Ein erster Versuch

Im vorigen Abschnitt wurde die Konstruktion eines Konfidenzintervalls $C(\underline{X})$ mittels einer Testfamilie $\{\Phi_{a_0, \sigma_0^2}(\underline{X}) : a_0 > 0, \sigma_0^2 > 0\}$ vorgestellt. Dies geschah durch Testen der Hypothesen

$$H_{a_0, \sigma_0^2}^0 : a = a_0, \sigma^2 \geq \sigma_0^2 \quad H_{a_0, \sigma_0^2}^1 : a > 0, \sigma^2 < \sigma_0^2, \quad a_0 > 0, \sigma_0^2 > 0.$$

Dabei wäre es möglich gewesen, $H_{a_0, \sigma_0^2}^1$ durch den Übergang zu

$$H_{a_0, \sigma_0^2}^{1*} : a = a_0, \sigma^2 < \sigma_0^2$$

weiter einzuschränken, ohne daß sich etwas geändert hätte.

Zur Hypothese $H_{a_0, \sigma_0^2}^0$ gegen $H_{a_0, \sigma_0^2}^{1*}$ läßt sich aber ein gleichmäßig bester Test $\Phi_{a_0, \sigma_0^2}^*(\underline{X})$ angeben. Aus einfachen Überlegungen ergibt sich

$$\Phi_{a_0, \sigma_0^2}^*(\underline{x}) = 1 \iff \frac{\bar{x}^2}{a_0} c_{a_0} < \sigma_0^2,$$

wobei $\frac{\bar{X}^2}{a_0}$ der Maximum-Likelihood-Schätzer der Varianz bei bekanntem Wert a_0 des Parameters a ist. Die Konstante c_{a_0} ergibt sich aus der Niveaubedingung

$$P_{a_0, \sigma_0^2} \left\{ \frac{\bar{X}^2}{a_0} c_{a_0} < \sigma_0^2 \right\} = \alpha,$$

was äquivalent zu

$$\mathcal{G}_{a_0 n} \left(\sqrt{\frac{a_0 n}{c_{a_0}}} \right) = \alpha$$

ist. Mit der Bezeichnung $\gamma_{a_0 n}$ für die Inverse von $\mathcal{G}_{a_0 n}$ lautet c_{a_0}

$$c_{a_0} = \frac{a_0 n}{\gamma_{a_0 n}^2(\alpha)}$$

und damit ergibt sich für das Konfidenzintervall $C^*(\underline{X})$ die Darstellung

$$C^*(\underline{X}) = \left\{ (a, \sigma^2) : \sigma^2 \leq \frac{n}{\gamma_{a n}^2(\alpha)} \bar{X}^2 \right\}$$

oder in der Notation des vorigen Abschnitts

$$C^*(\underline{X}) = \left\{ (a, \sigma^2) : \sigma^2 \leq c^*(a, \hat{a}_{ML}) \hat{\sigma}_{ML}^2 \right\}$$

mit $c^*(a, \hat{a}_{ML}) = \frac{n}{\gamma_{a n}^2(\alpha)} \hat{a}_{ML}$.

Projektion

Die Projektionsvorschrift $c(\hat{a}) = \sup_{a>0} c^*(a, \hat{a})$ liefert für alle $\hat{a} > 0$ das triviale Intervall $C(\underline{x}) = \{\sigma^2 : \sigma^2 \leq \infty\}$. Dies folgt aus der Anwendung der Markov-Ungleichung:

Da $\bar{X}$ eine nichtnegative Zufallsvariable mit existierendem Erwartungswert ist, gilt für beliebiges positives η :

$$\forall a > 0 \forall \sigma^2 > 0: P_{a, \sigma^2} \left\{ \bar{X} \geq \frac{\eta \sqrt{a \sigma^2}}{1 - \alpha} \right\} \leq E(\bar{X}) \frac{1 - \alpha}{\eta \sqrt{a \sigma^2}} = \frac{1 - \alpha}{\eta}$$

Dies ist äquivalent zu

$$\forall a > 0 \forall \sigma^2 > 0: \mathcal{G}_{a n, \frac{\sigma^2}{n}} \left(\frac{\eta \sqrt{a \sigma^2}}{1 - \alpha} \right) \geq 1 - \frac{1 - \alpha}{\eta}$$

und für $\eta > 1$ folgt

$$\forall a > 0: \mathcal{G}_{a n} \left(\frac{\eta \sqrt{a n}}{1 - \alpha} \right) > \alpha.$$

Da nach Definition

$$\forall a > 0: \mathcal{G}_{an}(\gamma_{an}(\alpha)) = \alpha$$

ist, folgt somit die für $\eta > 1$ gültige Ungleichung

$$\forall a > 0: \gamma_{an}(\alpha) < \frac{\eta\sqrt{an}}{1-\alpha}$$

und daraus

$$\lim_{a \rightarrow 0} \gamma_{an}^2(\alpha) = 0$$

sowie schließlich für $\hat{a} > 0$ die Behauptung

$$\sup_{a>0} c^*(a, \hat{a}) = \infty.$$

Das Konfidenzintervall besitzt die in Abbildung 4.7 angegebene Gestalt. Natürlich wäre es möglich, den Parameterraum durch Einführung einer unteren Schranke a_0 für a einzuschränken, so daß die Projektion kein triviales Intervall mehr ergäbe, doch ist das so konstruierte Konfidenzintervall immer noch sehr schlecht — die Funktionswerte von $\tilde{\alpha}$ betragen schon für Parameterwerte a , die nur geringfügig größer als a_0 sind, nahezu 0. Während die in Abschnitt 4.9 vorgeschlagene Funktion $c(\hat{a})$ antiton sein wird, steigt die in diesem Abschnitt konstruierte Funktion $c(\hat{a})$ linear in $\hat{a}$, so daß dieser Ansatz aufgegeben wird.

4.6. Ein bedingtes Konfidenzintervall

Der im letzten Abschnitt betrachtete Ansatz beruhte auf der Konstruktion eines gleichmäßig besten Konfidenzintervalls bei bekanntem Parameter a . Ein anderer Weg der Konstruktion besteht darin, a als unbekannt anzunehmen, und zu versuchen, den Parameter a durch die Durchführung eines bedingten Tests, wobei bezüglich der Zufallsvariablen $\hat{a}_{ML}$ bedingt wird, „in den Griff“ zu bekommen.

Zur Konstruktion

Für jedes $\hat{a} > 0$ soll $c(\hat{a})$ so gewählt werden, daß

$$\forall a > 0 \forall \sigma^2 > 0: P_{a,\sigma^2} \left\{ \sigma^2 \leq c(\hat{a}) \hat{\sigma}_{ML}^2 \mid \hat{a}_{ML} = \hat{a} \right\} \geq 1 - \alpha \quad (4.31)$$

erfüllt ist.

Aus der Einhaltung dieser Bedingung folgt unmittelbar

$$\forall a > 0 \forall \sigma^2 > 0:$$

$$\begin{aligned} P_{a,\sigma^2} \left\{ \sigma^2 \in C(\underline{X}) \right\} &= \int_0^\infty P_{a,\sigma^2} \left\{ \sigma^2 \leq c(\hat{a}) \hat{\sigma}_{ML}^2 \mid \hat{a}_{ML} = \hat{a} \right\} d\mathcal{A}_{a,n}(\hat{a}) \\ &\geq \int_0^\infty (1 - \alpha) d\mathcal{A}_{a,n}(\hat{a}) \\ &= 1 - \alpha \end{aligned}$$

die Einhaltung der Niveaubedingung.

Die Funktion c läßt sich nun punktweise aus der Bedingung (4.31) bestimmen.
Aus

$$\sigma^2 \leq c(\hat{a}) \hat{\sigma}^2 \iff \bar{X} \geq \sqrt{\sigma^2 \frac{\hat{a}}{c(\hat{a})}}$$

folgt aufgrund der Unabhängigkeit von $\bar{X}$ und $\hat{a}_{ML}$

$$P_{a, \sigma^2} \left\{ \sigma^2 \leq c(\hat{a}) \hat{\sigma}_{ML}^2 \mid \hat{a}_{ML} = \hat{a} \right\} = 1 - \mathcal{G}_{an, \frac{\sigma^2}{n}} \left(\sqrt{\sigma^2 \frac{\hat{a}}{c(\hat{a})}} \right) = 1 - \mathcal{G}_{an} \left(\sqrt{\frac{\hat{a}n}{c(\hat{a})}} \right).$$

Es ist also $c(\hat{a})$ so zu bestimmen, daß

$$\forall a > 0: \quad \mathcal{G}_{an} \left(\sqrt{\frac{\hat{a}n}{c(\hat{a})}} \right) \leq \alpha$$

erfüllt ist, oder mit der Inversen γ_{an} von $\mathcal{G}_{an}$ ausgedrückt

$$\forall a > 0: \quad c(\hat{a}) \geq \frac{n}{\gamma_{an}^2(\alpha)} \hat{a}.$$

Dieses ist, wie in Abschnitt 4.5 gezeigt wurde, für alle $\hat{a} > 0$ nur für $c(\hat{a}) = \infty$ erfüllt, das Konfidenzintervall ist im wesentlichen identisch mit dem Konfidenzintervall aus Abschnitt 4.5. Die Forderung, die Einhaltung der Niveaubedingung über die wesentlich stärkere Klasse von Bedingungen (4.31) erzwingen zu wollen, ist also nicht zweckmäßig.

4.7. Die Wahl der Startfunktion

Die in den beiden vorigen Abschnitten angegebenen Konfidenzintervalle waren offensichtlich nicht zur Konstruktion von $C(\underline{X})$ geeignet. Sie gehorchten zwar den in Abschnitt 4.4 aufgestellten Regeln, doch waren die Einschränkungen, die ihrer Konstruktion zugrundelagen, wohl zu stark. Es scheint insbesondere nicht zu gelingen, das fehlende Wissen über den Parameter a in einfacher Weise „auszugleichen“. In diesem Abschnitt sollen nun systematisch die Klasse aller nach den Überlegungen des Abschnitts 4.4 verbleibenden Konfidenzintervalle $C^*(\underline{X})$ und als Vorbereitung zu Abschnitt 4.9 eine gewisse Teilklasse daraus betrachtet werden. Zwar wird sich zeigen, daß auch in dieser Teilklasse die Zielvorstellung (4.7) nicht erfüllbar ist, doch es ergibt sich eine einfache Erweiterung des Ansatzes, die schließlich zum Ziel führt.

In Abschnitt 4.4 wurde die Dualität Konfidenzintervallkonstruktion-Testfamilie $\left\{ \Phi_{a_0, \sigma_0^2}(\underline{X}) : a_0 > 0, \sigma_0^2 > 0 \right\}$ vorgestellt, wobei $\Phi_{a_0, \sigma_0^2}(\underline{X})$ ein Test zu

$$H_{a_0, \sigma_0^2}^0 : a = a_0, \sigma^2 \geq \sigma_0^2 \quad H_{a_0, \sigma_0^2}^1 : a > 0, \sigma^2 < \sigma_0^2$$

ist. Aufgrund der Einseitigkeitsforderung (4.25) an $C^*(\underline{X})$, die ja eine zusätzlich Forderung an die Klasse $\left\{ \Phi_{a_0, \sigma_0^2}(\underline{X}) : a_0 > 0, \sigma_0^2 > 0 \right\}$ stellt, reicht es aus, sicherzustellen, daß

$$P_{a_0, \sigma_0^2} \left\{ \Phi_{a_0, \sigma_0^2}(\underline{X}) = 1 \right\} \leq \alpha \quad (4.32)$$

gilt, wodurch die zusammengesetzte Hypothese $H_{a_0, \sigma_0^2}^0$ zur einfachen Hypothese

$$H_{a_0, \sigma_0^2}^{0*} : a = a_0, \sigma^2 = \sigma_0^2$$

vereinfacht werden kann (dabei ist zu beachten, daß hierbei keine Einschränkung vorgenommen wird: Aus der Erfüllung der Niveaubedingung für die Verteilung aus $H_{a_0, \sigma_0^2}^{0*}$ folgt aufgrund der Einseitigkeitsforderung (4.25) automatisch die Einhaltung der Niveaubedingung für alle Verteilungen aus $H_{a_0, \sigma_0^2}^0$).

Pivotgrößen

Die Verteilungsfunktion der Zufallsvariablen $\hat{a}_{ML}$ soll als Zufallsvariable aufgefaßt mit U , die Verteilungsfunktion von $\bar{X}$ soll als Zufallsvariable aufgefaßt mit V bezeichnet werden:

$$U \stackrel{\text{def}}{=} \mathcal{A}_{a,n} \circ \hat{a}_{ML}, \quad V \stackrel{\text{def}}{=} \mathcal{G}_{a_n, \frac{\sigma^2}{n}} \circ \bar{X}$$

Mit $\bar{X}$ und $\hat{a}_{ML}$ sind auch U und V unabhängige Zufallsvariable, so daß gilt:

Sind $X_1, \dots, X_n$ unabhängige, identisch $G(a, \sigma^2)$ -verteilte Zufallsvariable, so sind $U = \mathcal{A}_{a,n} \circ \hat{a}_{ML}$ und $V = \mathcal{G}_{a_n, \frac{\sigma^2}{n}} \circ \bar{X}$ unabhängige, identisch auf dem Intervall $(0,1)$ gleichverteilte Zufallsvariable.

Die Gestalt der kritischen Region im Raum der Zufallsvariablen (U, V)

Die kritische Region des Tests $\Phi_{a_0, \sigma_0^2}(\underline{X})$ kann bijektiv auf den Raum $(0,1) \times (0,1)$ der Realisationen der Zufallsvariablen (U, V) abgebildet werden, wobei der Bildung von (U, V) die Verteilung aus $H_{a_0, \sigma_0^2}^{0*}$ zugrundegelegt wird:

$$U = \mathcal{A}_{a_0, n} \circ \hat{a}_{ML}, \quad V = \mathcal{G}_{a_0 n, \frac{\sigma_0^2}{n}} \circ \bar{X}$$

Dieser Bereich soll mit K_{a_0, σ_0^2} bezeichnet werden. Die Niveaubedingung (4.32) vereinfacht sich zur Flächenbedingung:

$$\text{Fläche} \left(K_{a_0, \sigma_0^2} \right) \leq \alpha \quad (4.33)$$

Die in Abschnitt 4.4 an die Testfamilie $\left\{ \Phi_{a_0, \sigma_0^2}(\underline{X}) : a_0 > 0, \sigma_0^2 > 0 \right\}$ gestellten Forderungen führten zum Konfidenzintervall

$$C^*(\underline{x}) = \left\{ (a_0, \sigma_0^2) : \sigma_0^2 \leq c^*(a_0, \hat{a}) \hat{\sigma}^2 \right\}.$$

Aufgrund der Beziehung

$$C^*(\underline{x}) = \left\{ (a_0, \sigma_0^2) : \Phi_{a_0, \sigma_0^2}(\underline{x}) = 0 \right\}$$

läßt sich damit die kritische Region K_{a_0, σ_0^2} , wie aus den Umformungen

$$\begin{aligned} \Phi_{a_0, \sigma_0^2}(\underline{x}) = 1 &\iff \frac{\bar{x}^2}{\hat{a}} c^*(a_0, \hat{a}) < \sigma_0^2 \\ &\iff \frac{\bar{x}}{\sigma_0} \sqrt{n} < \sqrt{\frac{\hat{a}n}{c^*(a_0, \hat{a})}} \\ &\iff v < \mathcal{G}_{a_0 n} \left(\sqrt{\frac{n \mathcal{A}_{a_0, n}^{-1}(u)}{c^*(a_0, \mathcal{A}_{a_0, n}^{-1}(u))}} \right) \end{aligned} \quad (4.34)$$

ersichtlich ist, stets durch eine auf $(0,1)$ definierte Funktion $h_{a_0}^*$ in der folgenden Weise

$$K_{a_0, \sigma_0^2} = \{(u, v) : v < h_{a_0}^*(u)\}$$

darstellen, wobei die Flächenbedingung (4.33) zur Integralbedingung

$$\int_0^1 h_{a_0}^*(u) du \leq \alpha \quad (4.35)$$

wird. Diese Darstellung zeigt, daß die Gestalt von K_{a_0, σ_0^2} nicht von σ_0^2 abhängen darf (der Grund hierfür ist die Annahme 2 aus Abschnitt 4.2). Die Funktion $h_{a_0}^*$ kann dabei unter Berücksichtigung von (4.35) frei vorgegeben werden, die Vorgabe der Familie $\{h_{a_0}^* : a_0 > 0\}$ entspricht der in Abschnitt 4.4 erwähnten „Freiheit“ der beliebigen Vorgabe einer Testfamilie $\{\Phi_{a_0, 1}(\underline{X}) : a_0 > 0\}$.

Damit läßt sich die in Abschnitt 4.4 aufgestellte Behauptung zeigen, es sei leicht $C^*(\underline{X})$ so zu konstruieren, daß gilt:

$$\forall a > 0 \quad \forall \sigma^2 > 0 : \quad P_{a, \sigma^2} \{ (a, \sigma^2) \in C^*(\underline{X}) \} = 1 - \alpha. \quad (4.36)$$

Im Raum der Realisationen von (U, V) ist lediglich eine beliebige Funktionenschar $\{h_a^* : a > 0\}$ vorzugeben, für welche in (4.35) stets das Gleichheitszeichen steht.

Zum Unterschied der Familien $\{h_a^* : a > 0\}$ und $\{h_a : a > 0\}$

Die Funktionenschar $\{h_a^* : a > 0\}$ ist von der in Abschnitt 4.2 eingeführten Familie $\{h_a : a > 0\}$ zu unterscheiden. Die frei vorzugebende Familie $\{h_a^* : a > 0\}$ definiert über den Zusammenhang

$$h_a^*(u) = \mathcal{G}_{an} \left(\sqrt{\frac{n \mathcal{A}_{a, n}^{-1}(u)}{c^*(a, \mathcal{A}_{a, n}^{-1}(u))}} \right) \quad (4.37)$$

die Funktion $c^*(a, \hat{a})$ und damit $C^*(\underline{X})$. Nach der Projektion ergibt sich $C(\underline{X})$ und daraus die Familie $\{h_a : a > 0\}$, die im allgemeinen mit der Familie $\{h_a^* : a > 0\}$ nicht übereinstimmt. Während die Familie $\{h_a^* : a > 0\}$ leicht so vorgegeben werden kann, daß gilt

$$\forall a > 0 : \int_0^1 h_a^*(u) du = \alpha, \quad (4.38)$$

ist es — wie in Abschnitt 4.1 erläutert — noch nicht einmal klar, ob es überhaupt eine Schar $\{h_a^* : a > 0\}$ gibt, für welche sich eine Familie $\{h_a : a > 0\}$ ergibt, so daß gilt:

$$\forall a > 0 : \alpha = \tilde{\alpha}(a) = \int_0^1 h_a(u) du.$$

Die beiden Familien stimmen genau dann exakt überein, wenn, wie aus (4.28), (4.37) und (4.21) leicht zu ersehen ist, die Funktion $c^*(a, \hat{a})$ im Argument a konstant ist, was nach den Erläuterungen des Abschnitts 4.4 (vgl. Abb. 4.8) bedeutet, daß für kein $\hat{a} > 0$ ein Projektionsverlust anfällt. In diesem Fall würde aus der Einhaltung von (4.37) die Erfüllung der Idealvorstellung (4.5) folgen. Ziel der Überlegungen dieses Abschnitts wird es sein, zu klären, ob eine Familie $\{h_a^* : a > 0\}$ dergestalt konstruierbar ist, daß wenigstens im Stützstellenbereich $\hat{a} \in \hat{A}$ kein Projektionsverlust auftritt unter gleichzeitiger näherungsweiser Einhaltung von (4.38) im Bereich $a \in A$.

Die Schar $\{h_a^* : a > 0\}$ zu den Konfidenzintervallen aus 4.5 und 4.6

In den Abschnitten 4.5 und 4.6 wurde ein Konfidenzintervall konstruiert, für welches sich $c^*(a, \hat{a}) = \frac{n}{\gamma_{an}^2(\alpha)} \hat{a}$ ergab. In (4.37) eingesetzt bedeutet dies, daß die Vorgabe der konstanten Funktion

$$h_a^*(u) = \mathcal{G}_{an} \left(\sqrt{\frac{n \mathcal{A}_{a,n}^{-1}(u)}{\frac{n}{\gamma_{an}^2(\alpha)} \mathcal{A}_{a,n}^{-1}(u)}} \right) = \mathcal{G}_{an}(\gamma_{an}(\alpha)) = \alpha$$

zu diesem Konfidenzintervall geführt hätte, ein Unterfangen, welches zwar (4.38) und damit (4.32) erfüllt, aber — wie in Abschnitt 4.5 gezeigt wurde — eine Funktionenschar $\{h_a : a > 0\}$ definiert, für welche

$$\forall a > 0 : \int_0^1 h_a(u) du = 0$$

gilt.

Ein erster Ansatz

In diesem Teil soll untersucht werden, welche Resultate sich aus der freien Vorgabe einer einzigen Funktion h^*

$$\forall a > 0: \quad h_a^*(\cdot) \stackrel{\text{def}}{=} h^*(\cdot)$$

ergeben. Dieser Ansatz erscheint plausibel, da U und V Pivotgrößen sind, also den Parameterwert a bereits einbeziehen.

Ein Beispiel

Die Erläuterungen sollen am Beispiel der in Abbildung 4.9 gezeichneten Funktion h^* gegeben werden.

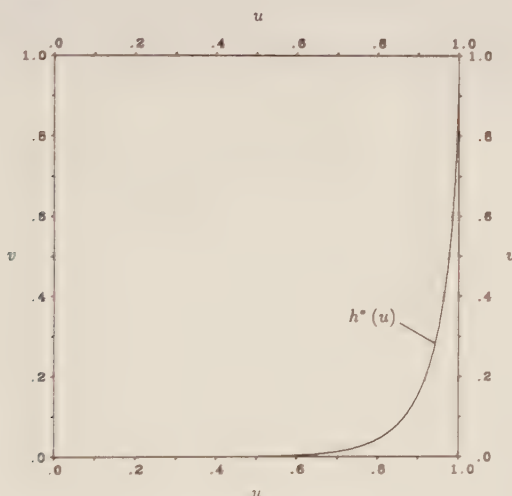


Abb. 4.9 Eine Beispielfunktion h^*

Anhand dieser Funktion, für welche

$$\int_0^1 h^*(u) \, du = 0,05$$

gilt, soll im folgenden die Konstruktion der Konfidenzintervalle $C^*(\underline{X})$ und $C(\underline{X})$ zum Stichprobenumfang $n = 50$ illustriert werden.

In einem ersten Schritt soll die Realisation von $C^*(\underline{X})$ für $(\hat{a}; \hat{\sigma}^2) = (4; 1)$ betrachtet werden. Nach (4.34) und (4.37) ist -- da $x = \sqrt{\hat{a}\hat{\sigma}^2} = 2$ gilt -- die

Nichtablehnung der Hypothese $H_{a_0, \sigma_0^2}^0$ gleichbedeutend mit

$$v = \mathcal{G}_{50 a_0} \left(\frac{2}{\sigma_0} \sqrt{50} \right) \geq h^* (\mathcal{A}_{a_0, 50} (4)) \quad (4.39)$$

Für $(a_0; \sigma_0^2) = (3; 1,5)$ ergibt sich $v = 0,246 > 0,148 = h^*(0,897)$ womit $(a_0; \sigma_0^2) = (3; 1,5)$ zur Realisation von $C^*(\underline{X})$ gehört, während sich für $(a_0; \sigma_0^2) = (3; 1,65)$, $v = 0,105 < 0,148 = h^*(0,897)$ und für $(a_0; \sigma_0^2) = (3; 1,593)$, $v = 0,148 = h^*(0,897)$ ergeben.

Aus (4.39) ist ersichtlich, daß $(a_0; \sigma_0^2) = (3; \sigma_0^2)$ genau dann in $C^*(\underline{x})$ liegt, wenn $\sigma_0^2 \leq 1,593$ gilt. In Abbildung 4.10 sind die obigen Tupel (u, v) eingetragen. Zudem sind noch die zu den Werten $(a_0; \sigma_0^2) = (2,75; 1,5)$, $(2,75; 1,65)$ und $(2,75; 1,593)$ gehörenden Paare (u, v) , für die $u = 0,953$ gilt, eingezeichnet. Dabei ist zu erkennen, daß das zu $(a_0; \sigma_0^2) = (2,75; 1,593)$ gehörende Paar $(u, v) = (0,953; 0,310)$ knapp unterhalb des Graphen von h^* liegt, so daß $(a_0; \sigma_0^2) = (2,75; 1,593)$ nicht in $C^*(\underline{x})$ liegt. $((2,75; \sigma_0^2)$ liegt genau dann in $C^*(\underline{x})$, wenn $\sigma_0^2 \leq 1,575$ erfüllt ist).

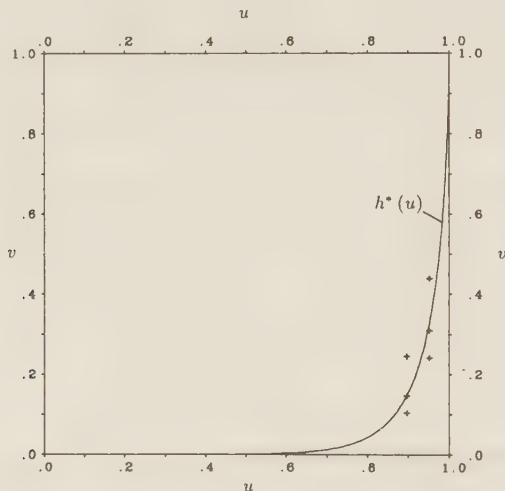


Abb. 4.10 Einige Realisationen von U, V

Wird a_0 weiter variiert, so ergibt sich aus der Projektionsvorschrift schließlich für $(\hat{a}; \hat{\sigma}^2) = (4; 1)$:

$$C(\underline{x}) = \{\sigma^2 : \sigma^2 \leq 1,646\}.$$

$c^*(a, \hat{a})$ besitzt also nicht die durch die Abb. 4.8 vorgegebene Gestalt, für $\hat{a} = 4$ tritt also ein Projektionsverlust auf.

Bevor das Beispiel für den Fall $\hat{a} \neq 4$ fortgesetzt wird, soll eine wichtige Funktionenschar eingeführt werden.

Die Familie der ξ -Funktionen

Im obigen Beispiel wurde für eine feste Realisation und für festes σ_0^2 der Wert a_0 variiert. Durchläuft a_0 den Bereich der positiven reellen Zahlen, so ergibt sich die folgende Familie von Tupeln

$$\left\{ (u, v) : u = \mathcal{A}_{a_0, 50}(4), v = \mathcal{G}_{50 a_0} \left(\frac{2}{\sigma_0} \sqrt{50} \right) : a_0 > 0 \right\}.$$

Diese Familie definiert eine Funktion mit Argument u , deren Graphen für $\sigma_0^2 = 1,5; 1,593$ und $1,65$ in Abbildung 4.11 mit $\xi_{1,5}$, $\xi_{1,593}$ und $\xi_{1,65}$ bezeichnet werden (eine genauere Untersuchung der Eigenschaften dieser Funktionen findet sich im Anschluß an diese Erläuterungen).

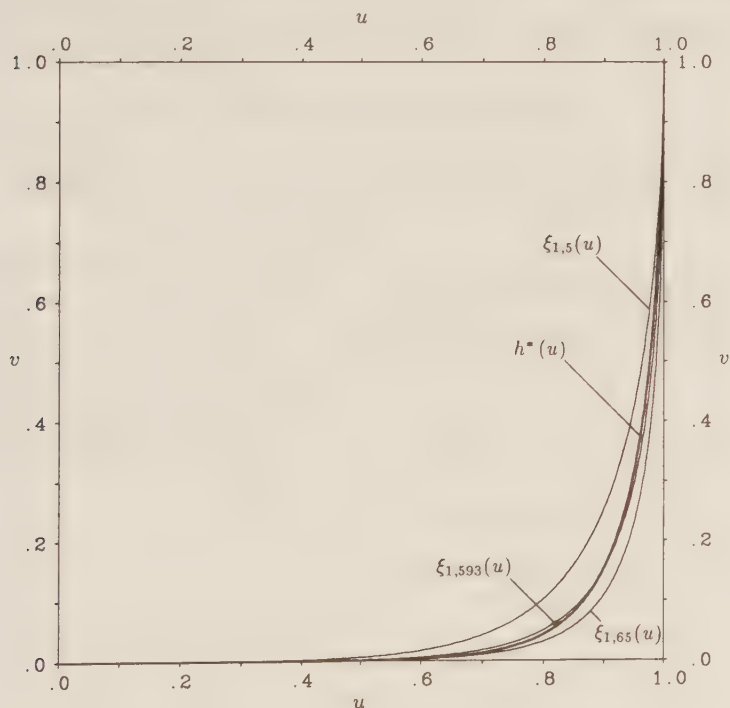


Abb. 4.11 Die Funktionen $\xi_{1,5}$, $\xi_{1,593}$ und $\xi_{1,65}$

Da jedem Punkt auf einem solchen Graphen $\xi_{\sigma_0^2}$ eine Realisation der Testgröße (U, V) zu einer Hypothese aus der Menge $\left\{H_{a_0, \sigma_0^2}^0 : a_0 > 0\right\}$ entspricht, wird σ_0^2 genau dann in das Konfidenzintervall $C(\underline{x})$ aufgenommen, wenn $\xi_{\sigma_0^2}$ an mindestens einer Stelle oberhalb des Graphen von h^* liegt. Falls $\xi_{\sigma_0^2}$ vollständig unterhalb von h^* liegt, wird jede Hypothese aus der obigen Menge abgelehnt, und σ_0^2 wird entsprechend der Projektionsvorschrift nicht in $C(\underline{x})$ aufgenommen. Im obigen Beispiel schneidet die Funktion $\xi_{1,593}$ die Funktion h^* an der Stelle $u = 0,897$, so daß $\sigma_0^2 = 1,593$ ins Konfidenzintervall $C(\underline{x})$ fällt. Die Funktion $\xi_{1,65}$ liegt dagegen im gesamten Bereich unterhalb von h^* , $\sigma_0^2 = 1,65$ wird nicht von $C(\underline{x})$ überdeckt. Im folgenden sollen einige Eigenschaften der ξ -Funktionen hergeleitet werden, wobei die Parametrisierung geändert werden soll, um zum Ausdruck zu bringen, daß sowohl der Wert des Parameters σ^2 als auch die Realisation $(\hat{a}, \hat{\sigma}^2)$ fest ist, und nur der Wert von a variiert.

Die Familie $\left\{\xi_{\hat{a}, \frac{\sigma^2}{\hat{\sigma}^2}} : \hat{a} > 0, \frac{\sigma^2}{\hat{\sigma}^2} > 0\right\}$

Zu einer festen Realisation $(\hat{a}, \hat{\sigma}^2)$ und festem Parameter σ^2 werden

$$u = \mathcal{A}_{a,n}(\hat{a}) \quad \text{und} \quad v = \mathcal{G}_{a,n, \frac{\sigma^2}{\hat{\sigma}^2}}(\bar{x}) = \mathcal{G}_{an} \left(\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}} \sqrt{\hat{a}n} \right)$$

als Funktionen des Parameters a aufgefaßt, was durch die Bezeichnungweise

$$u = \tilde{\mathcal{A}}_{\hat{a},n}(a) \quad \text{und} \quad v = \tilde{\mathcal{G}}_{\sqrt{\frac{\sigma^2}{\hat{\sigma}^2}} \sqrt{\hat{a}n}}(an)$$

zum Ausdruck gebracht werden soll.

I Die für positive a definierte Funktion

$$u : a \mapsto \tilde{\mathcal{A}}_{\hat{a},n}(a), \quad \hat{a} > 0$$

ist — wie in Kapitel 2.8 gezeigt — eine bijektive, stetige und streng antitone Abbildung in den Raum $(0, 1)$ mit

$$\lim_{a \rightarrow 0} \tilde{\mathcal{A}}_{\hat{a},n}(a) = 1 \quad \text{und} \quad \lim_{a \rightarrow \infty} \tilde{\mathcal{A}}_{\hat{a},n}(a) = 0.$$

Ihre Umkehrabbildung soll durch das Symbol

$$\tilde{\mathcal{A}}_{\hat{a},n}^{-1}$$

bezeichnet werden.

II Die im Intervall $(0, 1)$ definierte Funktion

$$\xi_{\hat{a}, \frac{\sigma^2}{\hat{\sigma}^2}} : u \mapsto \tilde{\mathcal{G}}_{\sqrt{\frac{\sigma^2}{\hat{\sigma}^2}} \sqrt{\hat{a}n}} \left(n \tilde{\mathcal{A}}_{\hat{a},n}^{-1}(u) \right), \quad \hat{\sigma}^2 > 0, \hat{a} > 0, \sigma^2 > 0$$

ist eine stetige Funktion mit

$$\lim_{u \rightarrow 0} \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}}(u) = 0 \quad \text{und} \quad \lim_{u \rightarrow 1} \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}}(u) = 1. \quad (4.40)$$

Der Beweis der Grenzwertaussage folgt aus I und mit Hilfe der in Kapitel 2.9 angegebenen, für alle positiven x gültigen Grenzwerte

$$\lim_{a \rightarrow 0} \mathcal{G}_a(x) = 1 \quad \text{und} \quad \lim_{a \rightarrow \infty} \mathcal{G}_a(x) = 0.$$

III Für

$$0 < \frac{\sigma_1^2}{\hat{\sigma}_1^2} < \frac{\sigma_2^2}{\hat{\sigma}_2^2} < \infty$$

gilt für alle $u \in (0, 1)$

$$\xi_{\hat{a}, \sigma_1^2 / \hat{\sigma}_1^2}(u) > \xi_{\hat{a}, \sigma_2^2 / \hat{\sigma}_2^2}(u). \quad (4.41)$$

Mit $a = \tilde{\mathcal{A}}_{\hat{a}, n}^{-1}(u)$ gilt $\xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}} \circ u = \mathcal{G}_{an} \left(\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}} \sqrt{\hat{a}n} \right)$, woraus die Behauptung aufgrund der strengen Isotonie von $\mathcal{G}_{an}$ folgt.

IV Die für positive a definierte Funktion v ergibt sich durch Hintereinanderschalten der Funktionen u und $\xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}}$

$$v = \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}} \circ u = \tilde{\mathcal{G}} \sqrt{\frac{\hat{\sigma}^2}{\sigma^2}} \sqrt{\hat{a}n}(an)$$

also

$$v : a \mapsto u = \tilde{\mathcal{A}}_{\hat{a}, n}(a) \mapsto \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}}(u) = \tilde{\mathcal{G}} \sqrt{\frac{\hat{\sigma}^2}{\sigma^2}} \sqrt{\hat{a}n} \left(n \tilde{\mathcal{A}}_{\hat{a}, n}^{-1}(u) \right). \quad (4.42)$$

Zur Konstruktion von $C(\underline{X})$ mit Hilfe der Familie $\left\{ \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}} : \hat{a} > 0, \frac{\sigma^2}{\sigma_0^2} > 0 \right\}$

Wie anhand des Beispiels demonstriert wurde, kann mit Hilfe der Familie $\left\{ \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}} : \hat{a} > 0, \frac{\sigma^2}{\sigma_0^2} > 0 \right\}$ die Realisation des Konfidenzintervalls $C(\underline{X})$ direkt bestimmt werden, also ohne den Umweg über die Bestimmung der Funktion $c^*(a, \hat{a})$ und der anschließenden Projektion:

Für eine gegebene Realisation $(\hat{a}, \hat{\sigma}^2)$ muß σ_0^2 so bestimmt werden, daß für alle $\sigma^2 > \sigma_0^2$ gilt

$$\forall u \in (0, 1) : \quad \xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}}(u) < h^*(u)$$

und für $\sigma^2 = \sigma_0^2$ muß mindestens ein $u \in (0, 1)$ existieren, für welches

$$\xi_{\hat{a}, \sigma_0^2 / \hat{\sigma}^2}(u) = h^*(u)$$

erfüllt ist. Für $\sigma^2 < \sigma_0^2$ existiert damit aufgrund der Monotoniebeziehung (4.41) mindestens ein $u \in (0, 1)$, für welches

$$\xi_{\hat{a}, \sigma_0^2 / \hat{\sigma}^2}(u) > h^*(u)$$

gilt. Die Realisation von $C(\underline{X})$ lautet damit

$$C(\underline{x}) = \left\{ \sigma^2 : \sigma^2 \leq \sigma_0^2 \right\},$$

was gleichbedeutend mit $\sigma_0^2 = c(\hat{a}) \hat{\sigma}^2$ und $\frac{\sigma_0^2}{\hat{\sigma}^2} = c(\hat{a})$ ist.

Zur Berechnung von c

Die Betrachtung der Familie $\left\{ \xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}} : \hat{a} > 0, \frac{\sigma^2}{\sigma^2} > 0 \right\}$ erlaubt also eine punktweise, direkte Bestimmung der Funktion c , die zudem aufgrund der Monotonieeigenschaft (4.41) auch aus numerischer Sicht noch sehr einfach gelingt. Die Konstruktion des Konfidenzintervalls wird daher im folgenden stets mit Hilfe dieser Familie durchgeführt.

Einige Überlegungen zur Gestalt der Funktion h^*

Aus den unter II erwähnten Eigenschaften der Funktion $\xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}}$ ergeben sich einige einfache Forderungen, die die Funktion h^* erfüllen sollte — aber natürlich nicht unbedingt erfüllen muß. Insbesondere sollten $\lim_{u \rightarrow 0} h^*(u) = 0$ und $\lim_{u \rightarrow 1} h^*(u) = 1$ gelten, da sich sonst — wie im Beispiel aus den Abschnitten 4.5 und 4.6 geschehen — $c(\hat{a}) = \infty$ ergibt. Streng genommen wurden sowohl der Quotient $\frac{\sigma^2}{\sigma^2} = \infty$ als auch $\frac{\sigma^2}{\sigma^2} = 0$ in der Definition II ausgeschlossen, die nötige Erweiterung ist aber durch Betrachtung der Grenzübergänge $\frac{\sigma^2}{\sigma^2} \rightarrow \infty$ bzw. $\frac{\sigma^2}{\sigma^2} \rightarrow 0$ augenscheinlich: Für positives $\hat{a}$ und $u \in (0, 1)$ ergeben sich $\xi_{\hat{a}, 0}(u) \stackrel{\text{def}}{=} 1$ und $\xi_{\hat{a}, \infty}(u) \stackrel{\text{def}}{=} 0$.

Die Funktion h^* sollte zudem im Intervall $(0, 1)$ stets positiv sein, sonst folgt wiederum $c(\hat{a}) = \infty$ (für die in Abb. 4.9 gezeichnete Funktion h^* gilt dieses, auch wenn dies aus Gründen der Zeichengenauigkeit nicht zu erkennen ist).

Projektionsverlust

Aus den bisherigen Erläuterungen ergibt sich der folgende Sachverhalt:

Bei der Projektion des Konfidenzintervalls $C^*(\underline{X})$ auf $C(\underline{X})$ gibt es genau dann zu einem festen Wert $\hat{a}$ keinen Projektionsverlust, wenn es einen Quotienten $c(\hat{a}) = \frac{\sigma^2}{\sigma^2}$ gibt, für den gilt:

$$\forall u \in (0, 1): \quad \xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}}(u) = h^*(u). \quad (4.43)$$

Gibt es einen Wert des Quotienten $\frac{\sigma^2}{\sigma^2}$, für welchen die zugehörige Funktion $\xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}}$ die Funktion h^* schneidet, so besagt die Monotonieeigenschaft (4.41), daß für das entsprechende $\hat{a}$ Projektionsverlust anfällt (s. obiges Beispiel: $\xi_{4,1,593}$ schneidet h^*). Eine Überprüfungsregel lautet:

Bestimme $\frac{\sigma^2}{\sigma^2}$, so daß gilt:

$$\int_0^1 \xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}}(u) du = \int_0^1 h^*(u) du \quad (4.44)$$

Ist die so bestimmte Funktion $\xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}}$ identisch mit h^* , so gilt $c(\hat{a}) = \frac{\sigma^2}{\sigma^2}$ und es fällt kein Projektionsverlust an. Ist die Identität (4.43) nicht erfüllt, so

entsteht ein Projektionsverlust und es gilt

$$c(\hat{a}) > \frac{\sigma^2}{\hat{\sigma}^2}.$$

In Abb. 4.12 sind einige Vertreter aus der Schar der Funktionen gezeichnet, die sich aus der Anwendung von (4.44) ergeben (im Beispiel gilt dabei $\int_0^1 h^*(u) du = 0,05$).

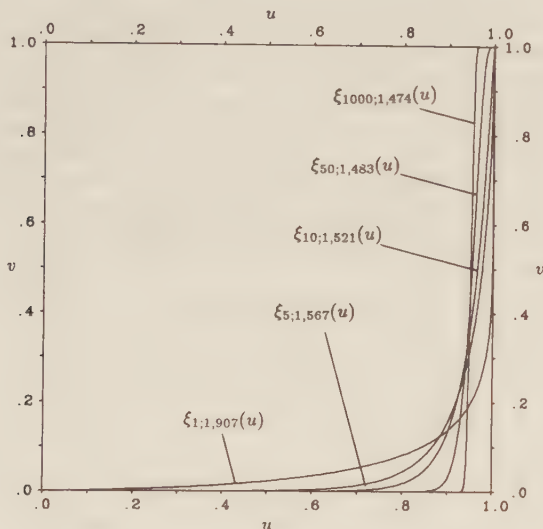


Abb. 4.12 Die Anwendung der Regel (4.44)

Da sich für unterschiedliche Werte von $\hat{a}$ offenkundig unterschiedliche Funktionen $\xi_{\hat{a}, \frac{\sigma^2}{\hat{\sigma}^2}}$ ergeben, scheint also der Versuch der Wahl einer einzigen Funktion h^* nicht geeignet zu sein, für alle $\hat{a} \in \hat{A}$ einen Projektionsverlust zu vermeiden. Eingezeichnet ist auch die Funktion $\xi_{5;1,567}$, die mit der Funktion h^* übereinstimmt. h^* wurde also so gewählt, daß für $\hat{a} = 5$ kein Projektionsverlust anfällt, für Werte von $\hat{a}$, die in der Nähe von 5 liegen, ist der Projektionsverlust — wie auch immer gemessen — relativ gering. Anhand der Gestalt der verschiedenen ξ -Funktionen kann in einfacher Weise die Güte des Verfahrens beurteilt werden, im Beispiel wird die Funktion $\tilde{a}$ im Bereich sehr kleiner und sehr großer Werte von a deutlich unter 0,05 liegen, in der Nähe des Wertes $a = 5$ relativ nahe am Wert 0,05 (aber immer noch unter 0,05, da es ja nur für die Beobachtung $\hat{a} = 5$ gelungen ist einen Projektionsverlust zu vermeiden). Die Abb. 4.13 zeigt die Funktion $\tilde{a}$, in Abb. 4.14 sind die für $\hat{a} \in \hat{A}$ nach der Regel (4.44) berechneten Quotienten $\frac{\sigma^2}{\hat{\sigma}^2}$

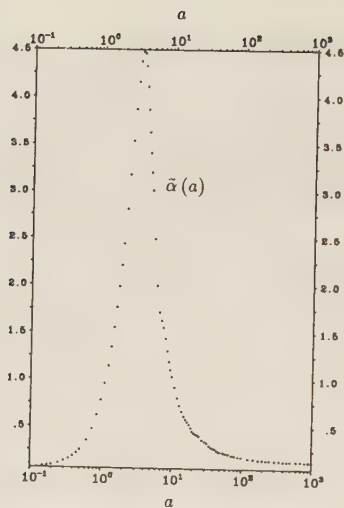


Abb. 4.13 Die Funktion $\tilde{\alpha}$

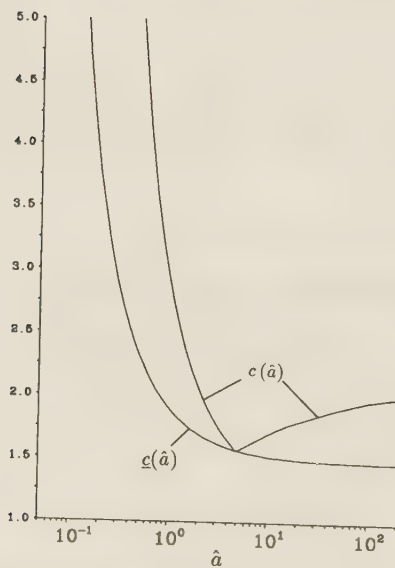


Abb. 4.14 Die Funktion $c(\hat{a})$ und ihre Abschätzung $\underline{c}(\hat{a})$ nach Regel (4.44)

eingetragen (und durch die untere, mit $\underline{c}$ bezeichnete, Kurve verbunden). Diese Punkte liefern den obigen Erläuterungen zufolge eine Abschätzung der sich aus h^* ergebenden Funktion $c(\hat{a})$ nach unten, $c(\hat{a})$ selbst ist ebenfalls eingezeichnet.

Die Wahl der Startfunktion $c^{(1)}$

In Abb. 4.15 ist „testweise“ diejenige Funktion $\tilde{\alpha}$ gezeichnet, die sich ergibt, wenn anstelle der Funktion $c(\hat{a})$ ihre Abschätzung gemäß Regel (4.44) zur Konstruktion von $C(\underline{X})$ herangezogen wird. Diese Funktion stellt in gewissem Sinne ein Ideal dar: Für jeden Wert $\hat{a} \in \hat{A}$ wird — natürlich unerlaubterweise — diejenige Funktion h^* zur Konstruktion herangezogen, die in dem gleichen Sinne für dieses $\hat{a}$ ideal ist, wie es im Beispiel die in Abb. 4.9 gezeichnete Funktion h^* für die Beobachtung $\hat{a} = 5$ ist. Wären die gemäß Regel (4.44) konstruierten ξ -Funktionen zusammengefallen, so hätte es, entsprechend der Aussage (4.43), eine Funktion h^* gegeben, die $\underline{c}$ geliefert hätte und für welche für kein $\hat{a} \in \hat{A}$ ein Projektionsverlust aufgetreten wäre. Falls dieses sogar für alle $\hat{a} > 0$ gegolten hätte, wäre damit sogar das Idealziel (4.5) erreicht worden. Da dieses aber nicht der Fall war, ergibt sich, daß dieses auf der „Idealvorstellung“ basierende Konfidenzintervall eine Funktion $\tilde{\alpha}$ besitzt, die oberhalb der Konstanten $\alpha = 0,05$ liegt, das Konfidenzintervall ist also zu „kurz“, wobei die Differenz $\tilde{\alpha}(a) - \alpha$ aber — wie in Abb. 4.15 zu erkennen ist — relativ gering ist.

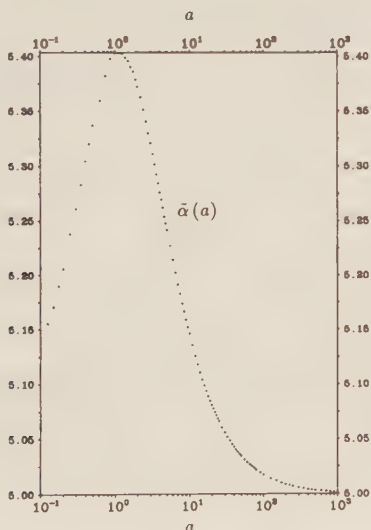


Abb. 4.15 Die $\tilde{\alpha}$ -Funktion des „Ideals“

Dieses gemäß der Regel (4.44) konstruierte Konfidenzintervall scheint damit ein geeigneter Anwärter für die in Abschnitt 4.3 eingeführte Startfunktion $c^{(1)}$ zu

sein. Daher soll seine Konstruktionsvorschrift — die oben als illegal bezeichnet wurde — durch die Angabe einer geeigneten Funktion h^* oder einer Familie $\{h_a^* : a > 0\}$ „legalisiert“ werden. Zum Konfidenzniveau $1 - \alpha$ ist dieses natürlich nicht möglich, zum Konfidenzniveau $1 - \sup_{a>0} \hat{\alpha}(a)$ muß es nach den zu Beginn dieses Abschnitts gegebenen Erläuterungen eine solche Familie geben (es stellte sich heraus, daß sich jedes Konfidenzintervall der Gestalt $C(\underline{x}) = \{\sigma^2 : \sigma^2 \leq c(\hat{a}) \hat{\sigma}^2\}$ durch eine Familie $\{h_a^* : a > 0\}$ darstellen läßt). Durch $\underline{c}$, das im folgenden bereits als Startfunktion $c^{(1)}$ bezeichnet wird, läßt sich z.B. die folgende Funktion $c^*(a, \hat{a})$ definieren:

$$\forall a > 0 \forall \hat{a} > 0 : c^*(a, \hat{a}) \stackrel{\text{def}}{=} c^{(1)}(\hat{a}).$$

Aus (4.37) ergibt sich damit direkt eine Familie $\{h_a^* : a > 0\}$, die noch dazu, wie unter dem Stichwort „Zum Unterschied der Familien $\{h_a^* : a > 0\}$ und $\{h_a : a > 0\}$ “ dargelegt, mit der Familie $\{h_a : a > 0\}$ zusammenfällt. Eine andere Möglichkeit bietet sich durch die Wahl der oberen Einhüllenden der Schar $\{\xi_{\hat{a}, c^{(1)}(\hat{a})} : \hat{a} > 0\}$ als Funktion h^* , wobei allerdings zu untersuchen bliebe, ob jede der Funktionen aus dieser Schar in wenigstens einem Punkt mit der so definierten Funktion h^* zusammenfällt.

Die obige Herleitung einer Familie $\{h_a^* : a > 0\}$ und damit eines Konfidenzintervalls $C^*(\underline{X})$ aus einem gegebenen Konfidenzintervall $C(\underline{X})$ stellt nun aber exakt das Gegenteil dessen dar, was mit der Einführung des ursprünglich gar nicht interessierenden Konfidenzintervalls $C^*(\underline{X})$ in Abschnitt 4.4 beabsichtigt war, nämlich die Konstruktion von $C(\underline{X})$ aus $C^*(\underline{X})$. Dieses Verfahren führte zur Einführung der Familie $\{\xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}} : \hat{a} > 0, \frac{\sigma^2}{\sigma_0^2} > 0\}$ und zur Regel (4.44). Die Funktion $\xi_{\hat{a}, \frac{\sigma^2}{\sigma_0^2}}$ faßt die Tests aller Hypothesen $H_{a_0, \sigma_0^2}^0 : a = a_0, \sigma^2 \geq \sigma_0^2$ aus der Familie $\{H_{a_0, \sigma_0^2}^0 : a_0 > 0\}$ zusammen und erweist sich mithin als geeignete zufällige Testfunktion der in Abschnitt 4.2 betrachteten zusammengesetzten Hypothese $H_{\sigma_0^2}^0 : \sigma^2 \geq \sigma_0^2, a > 0$ gegen $H_{\sigma_0^2}^1 : \sigma^2 < \sigma_0^2, a > 0$. Damit ist das in Abschnitt 4.4 verfolgte Ziel erreicht, im weiteren Verlauf wird nicht mehr vom Konfidenzintervall $C^*(\underline{X})$ die Rede sein. Bei Bedarf kann — wie oben demonstriert — jederzeit ein solches gemeinsames Konfidenzintervall angegeben werden.

Bevor im nächsten Abschnitt einige Eigenschaften des oben beschriebenen Verfahrens zu Konstruktion der Startfunktion $c^{(1)}$ diskutiert werden, soll der erste Schritt zusammenfassend noch einmal angegeben werden:

Die Startfunktion $c^{(1)}$ wird an den Stützstellen $\hat{a} \in \hat{A}$ so bestimmt, daß

$$\forall \hat{a} \in \hat{A} : \int_0^1 \xi_{\hat{a}, c^{(1)}(\hat{a})}(u) du = \alpha \quad (4.45)$$

erfüllt ist.

4.8. Eigenschaften der Startfunktion $c^{(1)}$

Vergleich der Familien $\{h_a : a > 0\}$ und $\{\xi_{\hat{a}, c(\hat{a})} : \hat{a} \in \hat{A}\}$

Aus den Definitionen der Familien $\{h_a : a > 0\}$ und $\{\xi_{\hat{a}, c(\hat{a})} : \hat{a} \in \hat{A}\}$ ergeben sich Beziehungen, die in diesem Abschnitt aufgezeigt werden sollen. Diese Eigenschaften erklären, warum die zur Startfunktion $c^{(1)}$ gehörende Funktion $1 - \tilde{\alpha}^{(1)}$ der Überdeckungswahrscheinlichkeiten nur so geringfügig vom Konfidenzniveau abweicht. In diesem Abschnitt soll der Index der Funktion $c^{(1)}$ der Kürze halber häufig fortgelassen werden.

In Abschnitt 4.1 wurde für die Funktionenschar $\{h_a : a > 0\}$, die über

$$P_{a, \sigma^2} \{ \sigma^2 \in C(\underline{X}) \} = 1 - \tilde{\alpha}(a) = 1 - \int_0^1 h_a(u) du \quad (4.46)$$

die Überdeckungswahrscheinlichkeiten definiert, für $\hat{a} > 0$ die zweistufige Beziehung

$$h_a \circ u : \hat{a} \mapsto u = \mathcal{A}_{a,n}(\hat{a}) \mapsto h_a(u) = \mathcal{G}_{an} \left(\sqrt{\frac{n\mathcal{A}_{a,n}^{-1}(u)}{c(\mathcal{A}_{a,n}^{-1}(u))}} \right) \quad (4.47)$$

angegeben. Für die Familie $\{\xi_{\hat{a}, c(\hat{a})} : \hat{a} \in \hat{A}\}$ gilt entsprechend

$$\xi_{\hat{a}, c(\hat{a})} \circ u : a \mapsto u = \tilde{\mathcal{A}}_{\hat{a}, n}(a) \mapsto \xi_{\hat{a}, c(\hat{a})}(u) = \tilde{\mathcal{G}}_{\sqrt{\frac{\hat{a}n}{c(\hat{a})}}} \left(n\tilde{\mathcal{A}}_{\hat{a}, n}^{-1}(u) \right) \quad (4.48)$$

Diese beiden Abbildungen lassen sich durch die Einführung der für positives $\hat{a}$ und a definierten Funktion

$$\eta : (\hat{a}, a) \mapsto \eta(\hat{a}, a) = \mathcal{G}_{an} \left(\sqrt{\frac{\hat{a}n}{c(\hat{a})}} \right)$$

als Schnitte

$$\begin{aligned} \forall \hat{a} > 0 : \quad \eta(\hat{a}, \cdot) &= \xi_{\hat{a}, c(\hat{a})} \circ u \\ \forall a > 0 : \quad \eta(\cdot, \hat{a}) &= h_a \circ u \end{aligned} \quad (4.49)$$

auffassen. Über (4.48) wird die Funktion c konstruiert:

$$1 - \alpha = 1 - \int_0^1 \xi_{\hat{a}, c(\hat{a})}(u) du \quad (4.50)$$

Es wird also bei festem Wert $\hat{a}$ im Endeffekt über den Parameter a integriert, während zur Erreichung der Zielvorstellung (4.7) c eigentlich so bestimmt werden müßte, daß sich bei der in (4.46) angegebenen Integration, bei der für festes a über $\hat{a}$ integriert wird, näherungsweise der Wert $1 - \alpha$ ergibt.

Zwei Annahmen

Es sollen nun zwei Annahmen für die Verteilungsfunktionen $\mathcal{A}_{a,n}$ und $\mathcal{G}_a$ angegeben werden, die — falls sie erfüllt wären — die Erreichung der Idealvorstellung (4.5) durch die Konstruktion der Startfunktion $c^{(1)}$ zur Folge hätten. Diese Annahmen sind zwar nicht erfüllt, stellen aber relativ gute Näherungen dar, so daß die gute Annäherung der Funktion $\hat{a}^{(1)}$ an die Konstante α verständlich wird.

1. Annahme

Die Verteilung der Zufallsvariablen $\frac{\hat{a}_{ML}}{a}$ hängt nur in relativ „geringem Maße“ vom Parameter a ab, in dem Sinne, daß sich die Verteilungsfunktion von $\frac{\hat{a}_{ML}}{a}$

$$P_a \left\{ \frac{\hat{a}_{ML}}{a} \leq z \right\} = P_a \{ \hat{a}_{ML} \leq az \} = \mathcal{A}_{a,n}(az)$$

für einen relativ großen Bereich von Werten a und z durch

$$\mathcal{A}_{a,n}(az) \approx \mathcal{A}_{1,n}(z)$$

approximieren läßt, wobei anstelle der Verteilungsfunktion $\mathcal{A}_{1,n}$ ohne Einschränkung der Allgemeinheit eine beliebige andere Verteilungsfunktion aus der Klasse der Verteilungsfunktionen $\{\mathcal{A}_{a,n} : a > 0\}$ gewählt werden könnte.

In Kapitel 2.6 wurde die Verteilung der Zufallsvariablen $\frac{\hat{a}_{ML}}{a}$ bereits betrachtet. Aus den dort angegebenen asymptotischen Verteilungen (Grenzübergänge $a \rightarrow 0$ und $a \rightarrow \infty$) ist ersichtlich, daß die obige Näherung durchaus brauchbar ist. Während in Kapitel 5 tatsächlich eine ähnliche Annahme für die Verteilungsfunktion der Zufallsvariablen Q^a (zum Zusammenhang zwischen Q^a und $\frac{\hat{a}_{ML}}{a}$ s. Kapitel 2.6) im Rahmen einer numerischen Näherung verwendet wird, soll an dieser Stelle keine Approximation vorgestellt werden, sondern es soll die „Qualität“ der Startfunktion c beurteilt werden. Die hier getroffene Annahme lautet:

Die Verteilung der Zufallsvariablen $\frac{\hat{a}_{ML}}{a}$ hängt vom Parameter a nicht mehr ab, $\frac{\hat{a}_{ML}}{a}$ ist also eine Pivotgröße.

Für die Familie der Verteilungsfunktionen $\{\mathcal{A}_{a,n} : a > 0\}$ gilt somit für alle positiven z :

$$\forall a > 0: \quad \mathcal{A}_{a,n}(z) = \mathcal{A}_{1,n}\left(\frac{z}{a}\right) \quad (4.51).$$

Die Umkehrfunktionen $\mathcal{A}_{a,n}^{-1}$ und $\tilde{\mathcal{A}}_{a,n}^{-1}$ ergeben sich mit (4.51) aus der Umkehrfunktion $\mathcal{A}_{1,n}^{-1}$:

$$\begin{aligned} \forall a > 0: \quad \mathcal{A}_{a,n}^{-1}(u) &= a \mathcal{A}_{1,n}^{-1}(u) \\ \forall \hat{a} > 0: \quad \tilde{\mathcal{A}}_{\hat{a},n}^{-1}(u) &= \frac{\hat{a}}{\mathcal{A}_{1,n}^{-1}(u)}. \end{aligned} \quad (4.52)$$

2. Annahme

Nach den in Kapitel 2.1 erläuterten Zusammenhängen ist die aus einer gammaverteilten Zufallsvariablen Z gebildete Transformation $\frac{Z}{E(Z)}$ ebenfalls gammaverteilt mit Gestaltparameter a und Skalierungsparameter $\frac{1}{a}$. Für diese Verteilung, die in Kapitel 5 zur Abschätzung der Funktion $\tilde{\alpha}$ herangezogen wird, liefert eine einfache Betrachtung ihrer Momente das Ergebnis, daß das r -te Moment $E\left(\left(\frac{Z}{E(Z)}\right)^r\right)$ für alle positiven a und alle $r \in [0, 1]$ beschränkt ist, während für die Zufallsvariablen Z , $\frac{Z}{b}$ und $\frac{Z}{\sigma}$ kein positives r existiert, für welches diese Eigenschaft, die den relativ geringen „Einfluß“ des Parameters a verdeutlicht, gilt. Somit kann auch die folgende Näherung der Verteilungsfunktion von $\frac{Z}{E(Z)}$

$$P_{a,\sigma^2} \left\{ \frac{Z}{E(Z)} \leq z \right\} = \mathcal{G}_a(z\sqrt{a}),$$

durch die Verteilungsfunktion $\mathcal{G}_1$

$$\mathcal{G}_a(z\sqrt{a}) \approx \mathcal{G}_1(z)$$

als erste grobe Näherung als brauchbar erachtet werden. Die zweite Annahme lautet auf die Verteilung von $\tilde{X}$ übertragen:

Die Verteilung der Zufallsvariablen $\frac{\tilde{X}}{E(\tilde{X})}$ hängt vom Parameter a nicht mehr ab, $\frac{\tilde{X}}{E(\tilde{X})}$ ist also eine Pivotgröße.

Es gilt damit für alle positiven z

$$\forall a > 0: \quad \mathcal{G}_{an}(z) = \mathcal{G}_1\left(\frac{z}{\sqrt{an}}\right), \quad (4.53)$$

bzw. in der Schreibweise mit vertauschtem Parameter und Argument für positives a

$$\forall z > 0: \quad \tilde{\mathcal{G}}_z(an) = \mathcal{G}_1\left(\frac{z}{\sqrt{an}}\right). \quad (4.54)$$

Die — in Wahrheit falschen — Formeln (4.52) und (4.54) können nun in (4.48) eingesetzt werden, die nach der Konstruktionsbedingung (4.45) für alle $\hat{a} \in \hat{A}$ zu erfüllende Beziehung lautet

$$\alpha = \int_0^1 \tilde{\mathcal{G}}_{\sqrt{\frac{\hat{a}n}{c(\hat{a})}}} \left(n \tilde{\mathcal{A}}_{\hat{a},n}^{-1}(u) \right) du = \int_0^1 \mathcal{G}_1 \left(\sqrt{\frac{\mathcal{A}_{1,n}^{-1}(u)}{c(\hat{a})}} \right).$$

Damit ergibt sich für alle $\hat{a} \in \hat{A}$ die gleiche Lösung, c ist eine konstante Funktion:

$$c(\cdot) \equiv \tilde{c}.$$

Unter Verwendung von (4.52) und (4.53) ergibt sich daraus für die Familie $\{h_a : a > 0\}$

$$h_a(u) = \mathcal{G}_{an} \left(\sqrt{\frac{n\mathcal{A}_{a,n}^{-1}(u)}{c(\mathcal{A}_{a,n}^{-1}(u))}} \right) = \mathcal{G}_1 \left(\sqrt{\frac{\mathcal{A}_{1,n}^{-1}(u)}{\tilde{c}}} \right),$$

was bedeutet, daß auch diese Familie nur aus einer einzigen Funktion besteht, die noch dazu mit der die Familie $\{\xi_{\hat{a},c(\hat{a})} : \hat{a} \in \hat{A}\}$ darstellenden Funktion identisch ist, so daß die Konstruktionsbedingung (4.45) — auf $\hat{a} > 0$ erweitert — sofort ein ähnliches Konfidenzintervall liefert.

Eine schwächere Annahme

Die Zeichnung der Funktionenschar $\{\xi_{\hat{a},c(\hat{a})} : \hat{a} \in \hat{A}\}$ des im letzten Abschnitt besprochenen Beispiels (s. Abb. 4.12) zeigt, daß die Funktionen sich doch noch recht deutlich unterscheiden, die obigen Pivotannahmen also relativ grob sind. Dies ist aber nicht gravierend. Es ist keineswegs notwendig, daß alle Funktionen aus $\{\xi_{\hat{a},c(\hat{a})} : \hat{a} > 0\}$ und $\{h_a : a > 0\}$ zusammenfallen. Wäre die schwächere Annahme der Existenz einer bijektiven Abbildung zwischen den Familien $\{\xi_{\hat{a},c(\hat{a})} : \hat{a} > 0\}$ und $\{h_a : a > 0\}$ erfüllt, so würde die Anwendung der Regel:

Bestimme c , so daß

$$\forall \hat{a} > 0 : \int_0^1 \xi_{\hat{a},c(\hat{a})}(u) du = \alpha$$

gilt,

ebenfalls die Erfüllung der Zielvorstellung (4.5) bedeuten. Natürlich existiert auch solch eine Abbildung nicht, doch (4.51) und (4.53) sind für nicht zu weit auseinanderliegende Werte a_0 und a_1 in guter Näherung erfüllt, d.h. für alle $z > 0$ gilt

$$\mathcal{A}_{a_0,n}(za_0) \approx \mathcal{A}_{a_1,n}(za_1)$$

und

$$\mathcal{G}_{a_0n}(z\sqrt{a_0n}) \approx \mathcal{G}_{a_1n}(z\sqrt{a_1n}),$$

woraus sich die in Tabelle 4.2 und Abbildung 4.16 demonstrierte Übereinstimmung der Familien begründen läßt. In Tab. 4.2 ist für einige Werte von $\hat{a}$

$$\int_0^1 |\xi_{\hat{a},c(\hat{a})}(u) - h_{a_0}(u)| du = \min_{a \in A} \int_0^1 |\xi_{\hat{a},c(\hat{a})}(u) - h_a(u)| du$$

angegeben. Diese Werte müßten — falls nicht nur der Stützstellenbereich A betrachtet würde — im Idealfall der Existenz einer solchen bijektiven Abbildung 0 betragen.

$\hat{a}$	a_0	$\int_0^1 \xi_{\hat{a}, c(\hat{a})}(u) - h_{a_0}(u) du$
0,10	0,10	0,0019
0,15	0,125	0,0020
0,25	0,20	0,0024
0,35	0,30	0,0028
0,50	0,40	0,0032
0,75	0,625	0,0037
1	0,75	0,0040
1,5	1,25	0,0041
2	1,625	0,0039
3	2,4	0,0036
4	3,2	0,0032
5	3,8	0,0028
7,5	5,5	0,0023
10	7,5	0,0019
15	11	0,0014
20	14	0,0012
25	18	0,00093
30	21	0,00080
40	27,5	0,00066
50	35	0,00049
75	50	0,00037
100	70	0,00028
250	175	0,00012
500	350	0,00008
1000	700	0,00006

Tab. 4.2 Zur Übereinstimmung der Familien $\{\xi_{\hat{a}, c(\hat{a})} : \hat{a} \in \hat{A}\}$ und $\{h_a : a \in A\}$

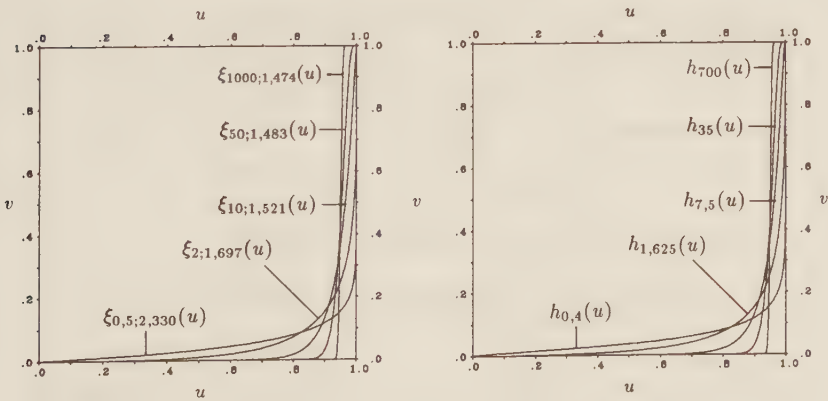


Abb. 4.16 Einige Funktionenpaare $(\xi_{\hat{a}, c(\hat{a})}, h_{a_0})$ aus Tab. 4.2

In Abb. 4.16 sind einige der in Tab. 4.2 angegebenen Paare $(\xi_{\hat{a}, c(\hat{a})}, h_a)$ gegenübergestellt. Es ist deutlich zu erkennen, daß die Funktionen sich nur geringfügig unterscheiden, die Näherung ist also sehr gut.

4.9. Der Iterationschritt von $c^{(i)}$ nach $c^{(i+1)}$

Bevor der Schritt der Berechnung von $c^{(i+1)}$ aus $c^{(i)}$ und $\tilde{\alpha}^{(i)}$ erläutert wird, soll noch kurz der in Abschnitt 4.3 erwähnte Interpolationsalgorithmus besprochen werden, der die nur im Stützstellenbereich $\hat{A}$ numerisch bestimmte Funktion c auf den gesamten Raum $\mathbb{R}^+$ fortsetzt.

Der Interpolationsalgorithmus

In Abb. 4.17 wird die Funktion $c^{(1)}$ für die Stichprobenumfänge $n = 10, 50, 250$ und $\alpha = 0,05$ an den Stützstellen gezeichnet.

Die Gestalt von $c^{(1)}$ erinnert in allen drei Fällen an eine Exponentialfunktion, allerdings als Funktion von $\frac{1}{\hat{a}}$ aufgefaßt. In Abb. 4.18 sind zur Überprüfung deshalb die drei Funktionen $\ln c^{(1)}$ des obigen Beispiels gegen $\frac{1}{\hat{a}}$ angetragen.

Die zwischen den Stützstellen der obigen Funktionen sicherlich gerechtfertigte lineare Interpolation führt zur folgenden Vorschrift:

Die Funktion c sei im Bereich $\hat{A} = \{\hat{a}_1, \dots, \hat{a}_{104}\}$ mit $\hat{a}_1 < \hat{a}_2 < \dots < \hat{a}_{103} < \hat{a}_{104}$ gegeben. Für $\hat{a}_1 < \hat{a} < \hat{a}_{104}$ wird $c(\hat{a})$ wie folgt definiert:

$$\forall \hat{a} \in (\hat{a}_j, \hat{a}_{j+1}): \quad c(\hat{a}) \stackrel{\text{def}}{=} \gamma_j \exp\left(\frac{\delta_j}{\hat{a}}\right), \quad j = 1, \dots, 103,$$

wobei γ_j und δ_j so bestimmt werden, daß

$$c(\hat{a}_j) = \gamma_j \exp\left(\frac{\delta_j}{\hat{a}_j}\right) \quad \text{und} \quad c(\hat{a}_{j+1}) = \gamma_j \exp\left(\frac{\delta_j}{\hat{a}_{j+1}}\right)$$

erfüllt sind. Für $\hat{a} < \hat{a}_1$ wird die Vorschrift für das Intervall $(\hat{a}_1, \hat{a}_2)$ übernommen:

$$\forall \hat{a} < \hat{a}_1: \quad c(\hat{a}) \stackrel{\text{def}}{=} \gamma_1 \exp\left(\frac{\delta_1}{\hat{a}}\right),$$

für $\hat{a} > \hat{a}_{104}$ wird

$$\forall \hat{a} > \hat{a}_{104}: \quad c(\hat{a}) \stackrel{\text{def}}{=} \gamma_{104} \exp\left(\frac{\delta_{104}}{\hat{a}}\right)$$

definiert, wobei γ_{104} und δ_{104} so bestimmt werden, daß

$$\gamma_{104} = \frac{n}{\mathcal{G}_{\frac{(n-1)}{2}, 2(n-1)}^{-1}(\alpha)} \quad \text{und} \quad c(\hat{a}_{104}) = \gamma_{104} \exp\left(\frac{\delta_{104}}{\hat{a}_{104}}\right)$$

erfüllt sind.

Die Begründung für die Wahl der Konstanten γ_{104} wird am Ende dieses Abschnitts gegeben, die Konstante $\mathcal{G}_{\frac{(n-1)}{2}, 2(n-1)}^{-1}(\alpha)$ bezeichnet das α -Quantil der χ^2 -Verteilung mit $n - 1$ Freiheitsgraden.

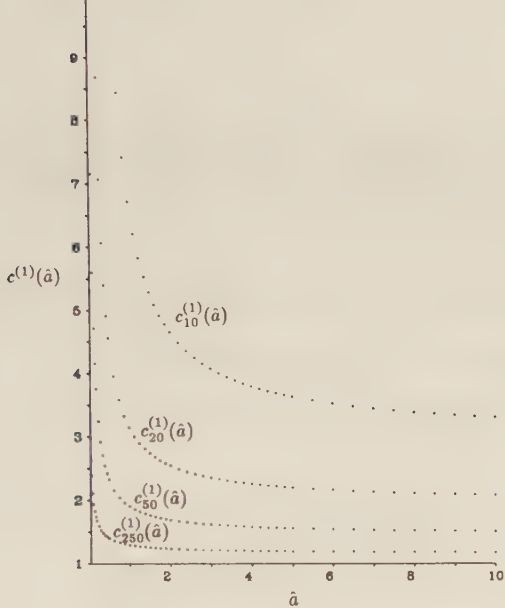


Abb. 4.17 Die Funktion $c^{(1)}$ für $n = 10, 50, 250$

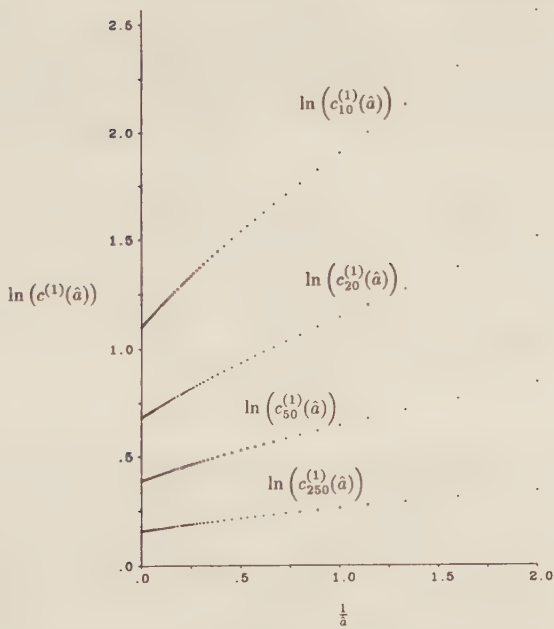


Abb. 4.18 Die Funktion $\ln(c^{(1)})$ für $n = 10, 50, 250$

Iterationsschritt

Die Konstruktion der Funktion $c^{(i+1)}$ ($i > 0$) wird wiederum über eine Integralbedingung an die Schar $\{\xi_{\hat{a}, c^{(i+1)}(\hat{a})} : \hat{a} \in \hat{A}\}$ erfolgen. Während $c^{(1)}$ über die Bedingung

$$\forall \hat{a} \in \hat{A} : \int_0^1 \xi_{\hat{a}, c^{(1)}(\hat{a})}(u) du = \alpha$$

konstruiert wird, soll $c^{(i+1)}$ ($i > 0$) über die Bedingung

$$\forall \hat{a} \in \hat{A} : \int_0^1 \xi_{\hat{a}, c^{(i+1)}(\hat{a})}(u) du = \beta^{(i+1)}(\hat{a}) \quad (4.55)$$

gebildet werden, wobei die Funktion $\beta^{(i+1)}$ aus $\beta^{(i)}$ und $\tilde{\alpha}^{(i)}$ berechnet wird. Die Konstruktionsvorschrift (4.55) schließt durch die Wahl von

$$\forall \hat{a} \in \hat{A} : \beta^{(1)}(\hat{a}) \stackrel{\text{def}}{=} \alpha$$

natürlich auch die Regel (4.45) ein.

In den Abbildungen 4.14 und 4.15 wurden die Funktionen $c^{(1)}$ und $\tilde{\alpha}^{(1)}$ für das Beispiel $n = 50$, $1 - \alpha = 0,95$ gezeichnet. Die Funktion $\tilde{\alpha}^{(1)}$ liegt im gesamten Bereich oberhalb der Konstanten α , so daß $c^{(2)}$ vollständig, d.h. für alle $\hat{a} \in \hat{A}$ oberhalb von $c^{(1)}$ liegen sollte. Aus der Monotoniebeziehung (4.41) folgt, daß dieses durch eine gleichmäßige Verringerung von $\beta^{(1)}$ erreicht wird, es ist $\beta^{(2)}$ also so zu wählen, daß

$$\forall \hat{a} \in \hat{A} : \beta^{(2)}(\hat{a}) < \beta^{(1)}(\hat{a}) = \alpha$$

gilt. Aufgrund der im letzten Abschnitt betrachteten Beziehung zwischen der Funktion $\beta^{(1)}$ und der Funktion $\tilde{\alpha}^{(1)}$ — die natürlich analog zwischen $\beta^{(2)}$ und $\tilde{\alpha}^{(2)}$ besteht — und da $\hat{a}_{ML}$ der Maximum-Likelihood-Schätzer von a ist, liegt es nahe, $\beta^{(2)}$ so zu wählen, daß der Quotient der Funktionen $\frac{\beta^{(2)}}{\beta^{(1)}} = \frac{\beta^{(2)}}{\alpha}$ gleich dem Quotienten der Funktionen $\frac{\alpha}{\tilde{\alpha}^{(1)}}$ ist, also

$$\forall \hat{a} \in \hat{A} : \beta^{(2)}(\hat{a}) \stackrel{\text{def}}{=} \beta^{(1)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(1)}(\hat{a})}. \quad (4.56)$$

Da die Stützstellenmenge A fast identisch zur Stützstellenmenge $\hat{A}$ ist, kann (4.56) direkt angewendet werden, lediglich an den sieben zu Beginn des Abschnitts 4.3 angegebenen Stützstellen ist $\tilde{\alpha}^{(1)}(\hat{a})$ nicht berechnet, so daß hier zusätzliche Überlegungen, die an späterer Stelle erläutert werden, nötig sind. Für kleine Werte von $\hat{a}$ fällt der Wert von $\beta^{(1)}$ — und damit der Wert von $c^{(1)}$ — stärker als für mittlere und große Werte von $\hat{a}$. Dieses erscheint geeignet, die entsprechend große Differenz von $\tilde{\alpha}^{(1)}$ und α im Bereich kleiner Werte von a zu korrigieren, während

für mittlere und große Werte von a die bessere Übereinstimmung der Funktion $\tilde{\alpha}^{(1)}$ mit der Konstanten α eine geringere Korrektur von $\beta^{(1)}$ im Bereich mittlerer und größerer Werte von $\hat{a}$ nötig macht. Die Vorschrift (4.56) bedeutet eine für alle $\hat{a} \in A$ vorgenommene Korrektur von $\beta^{(1)}(\hat{a})$ durch den Funktionswert von $\tilde{\alpha}^{(1)}$ an der Stelle $a = \hat{a}$, also eine punktweise Korrektur, die zu guten Ergebnissen führt, da $\hat{a}_{ML}$ ein „geeigneter“ Punktschätzer des Parameters a ist.

In Abb. 4.19 sind für das Beispiel $n = 50, 1 - \alpha = 0,95$ die sich auf diese Weise ergebenden Funktionen $\tilde{\alpha}^{(1)}, \tilde{\alpha}^{(2)}$ und $\tilde{\alpha}^{(3)}$ angegeben, wobei der Iterationsschritt der Bildung von $\beta^{(3)}$ aus $\beta^{(2)}$ analog

$$\forall \hat{a} \in \hat{A}: \quad \beta^{(3)}(\hat{a}) \stackrel{\text{def}}{=} \beta^{(2)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(2)}(\hat{a})}$$

lautet.

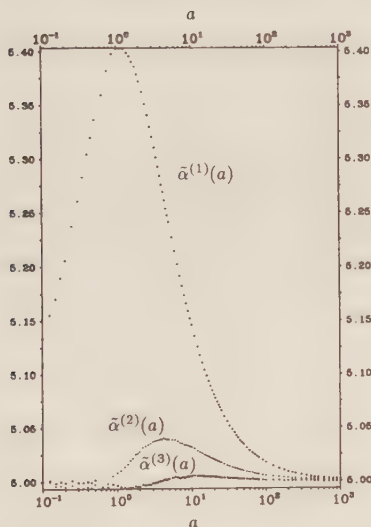


Abb. 4.19 Die Funktionen $\tilde{\alpha}^{(i)}$, $i = 1, 2, 3$ des Beispiels $n = 50, 1 - \alpha = 0,95$

Die Größenordnung der sich daraus ergebenden Korrektur der Funktion $c^{(1)}$ ist, verglichen mit der absoluten Größenordnung von $c^{(1)}$, sehr gering, so daß auf eine entsprechende Zeichnung verzichtet werden soll. Die maximale relative Differenz der Funktionen $c^{(1)}$ und $c^{(2)}$ beträgt lediglich $\max_{\hat{a} \in \hat{A}} \left| \frac{c^{(2)}(\hat{a}) - c^{(1)}(\hat{a})}{c^{(1)}(\hat{a})} \right| = 0,018$, für den Vergleich von $c^{(3)}$ und $c^{(2)}$ ergibt sich sogar $\max_{\hat{a} \in \hat{A}} \left| \frac{c^{(3)}(\hat{a}) - c^{(2)}(\hat{a})}{c^{(2)}(\hat{a})} \right| < 0,001$.

Da zur Korrektur von $c^{(i)}$ nur die Werte der Funktion $\tilde{\alpha}^{(i)}$ im Bereich $a \in \hat{A}$ herangezogen werden, würde die Wahl von $A = \hat{A}$ eine Interpolationsvorschrift zur Schätzung von $\tilde{\alpha}^{(i)}$ außerhalb des Bereichs A unnötig machen. Allerdings ist die Wahl von $A = \hat{A}$ problematisch, da der Funktionswert von $\tilde{\alpha}^{(i)}$ im Bereich kleiner Werte von a maßgeblich vom Verlauf der Funktion $c^{(i)}$ im Bereich kleiner $\hat{a}$ abhängt — in diesem Bereich wären dann allerdings zu wenige Stützstellen vorgegeben, und der Wert von $\tilde{\alpha}^{(i)}$ würde wesentlich vom Interpolationsalgorithmus bestimmt. Daher ist es sinnvoller — wie in Abschnitt 4.3 vorgenommen — den Stützstellenbereich $\hat{A}$ gegenüber dem Bereich A zu erweitern, und für $\hat{a} \notin \hat{A}$ die Regel (4.56) für $i \geq 1$ folgendermaßen zu modifizieren:

$$\begin{aligned} \forall \hat{a} < a_{\min} : \quad \beta^{(i+1)}(\hat{a}) &\stackrel{\text{def}}{=} \beta^{(i)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(i)}(a_{\min})} \\ \forall \hat{a} \in A : \quad \beta^{(i+1)}(\hat{a}) &\stackrel{\text{def}}{=} \beta^{(i)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(i)}(\hat{a})} \\ \forall \hat{a} > a_{\max} : \quad \beta^{(i+1)}(\hat{a}) &\stackrel{\text{def}}{=} \beta^{(i)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(i)}(a_{\max})} \end{aligned} \quad (4.57)$$

wobei mit $a_{\min}$ und $a_{\max}$ das minimale bzw. das maximale Element der Menge A bezeichnet wird.

Erwartungswertkorrektur

Bevor einige weitere Beispiele für die Funktionsweise der Iteration angegeben werden, soll noch eine einfache Modifikation der Vorschrift (4.57) vorgestellt werden, die für kleine Stichprobenumfänge die Erfüllung der Zielvorstellung (4.7) in einer geringeren Anzahl von Iterationsschritten zur Folge hat. Diese Korrektur ist eine einfache Biaskorrektur und verwendet die in Kapitel 2.6 angegebenen Beziehung

$$E(\hat{a}_{ML}) = \frac{n}{n-3}a + o(1) \quad \text{für } a \rightarrow \infty.$$

Die Korrektur verläuft analog wie unter (4.57) angegeben, anstelle von

$$\beta^{(i+1)}(\hat{a}) \stackrel{\text{def}}{=} \beta^{(i)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(i)}(\hat{a})}$$

wird lediglich

$$\beta^{(i+1)}(\hat{a}) \stackrel{\text{def}}{=} \beta^{(i)}(\hat{a}) \frac{\alpha}{\tilde{\alpha}^{(i)}\left(\frac{n}{n-3}\hat{a}\right)}$$

gesetzt. Selbstverständlich sollte bei dieser Vorgehensweise die Menge A entsprechend der Erwartungswertkorrektur modifiziert werden.

Einige numerische Vergleiche

In Tab. 4.3 werden für $\alpha = 0,01; 0,05; 0,10$, $n = 10, 20, 50, 250$ und $i = 1, \dots, 10$ die Differenzen

$$\min_{a \in A} 10^6 \left(\tilde{\alpha}^{(i)}(a) - \alpha \right), \quad \max_{a \in A} 10^6 \left(\tilde{\alpha}^{(i)}(a) - \alpha \right)$$

Table 1								
No.	1950		1951		1952		1953	
	Jan	Feb	Jan	Feb	Jan	Feb	Jan	Feb
1	100	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100	100

Table 2								
No.	1950		1951		1952		1953	
	Jan	Feb	Jan	Feb	Jan	Feb	Jan	Feb
1	100	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100	100

Table 3								
No.	1950		1951		1952		1953	
	Jan	Feb	Jan	Feb	Jan	Feb	Jan	Feb
1	100	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100	100

Table 4: Summary of data for 1950, 1951, 1952, and 1953.

angegeben. Zu diesen Werten ist anzumerken, daß die Konstruktionsbedingung (4.55) in der folgenden Weise abgeschwächt wurde:

Bestimme $c^{(i)}$, so daß

$$\forall \hat{a} \in \hat{A}: \quad \beta^{(i)}(\hat{a}) - 10^{-5} \leq \int_0^1 \xi_{\hat{a}, c^{(i)}(\hat{a})}(u) du \leq \beta^{(i)}(\hat{a}) + 10^{-5}$$

erfüllt ist.

Schon diese, aus numerischen Gründen zweckmäßige Festsetzung alleine verhindert, daß die Konstante α mit beliebig hoher Genauigkeit angenähert werden kann. Der Bereich, in dem der Funktionswert $c^{(i)}(\hat{a})$ schwanken kann, ist allerdings winzig, so daß diese Abschwächung keine praktische Bedeutung hat. Aufgrund dieser Schwankungsbreite ist klar, daß auch die Funktion $\tilde{\alpha}^{(i)}$ eine Schwankungsbreite um α von mindestens 10^{-5} aufweist, die tabellierten Minima und Maxima werden im allgemeinen dem Betrag nach größer als 10^{-5} (Tabelleneintrag: 10) sein. Es zeigt sich eine sehr schnelle gleichmäßige Annäherung der Funktion $\tilde{\alpha}^{(i)}$ an die Konstante α , lediglich für $\alpha = 0,01$ und $n = 10$ beträgt die betragsmäßig größte Abweichung in allen Fällen mehr als 0,0001. Aus der Tabelle läßt sich entnehmen, für welche Werte von ϵ die Zielvorstellung (4.7) erfüllbar ist.

Der Grenzübergang $\hat{a} \rightarrow \infty$

In Kapitel 2.7 wurde die für $a \rightarrow \infty$ gültige Grenzverteilung

$$\frac{n\hat{\sigma}_{ML}^2}{\sigma^2} \sim \chi_{n-1}^2 \quad \text{für } a \rightarrow \infty$$

hergeleitet. Mit der Bezeichnung $\mathcal{G}_{\frac{(n-1)}{2}, 2(n-1)}^{-1}(\alpha)$ für das α -Quantil der χ_{n-1}^2 -Verteilung ergibt sich für

$$c_\infty \stackrel{\text{def}}{=} \frac{n}{\mathcal{G}_{\frac{(n-1)}{2}, 2(n-1)}^{-1}(\alpha)}$$

unmittelbar

$$\lim_{a \rightarrow \infty} P_{a, \sigma^2} \left\{ \sigma^2 \leq c_\infty \hat{\sigma}_{ML}^2 \right\} = 1 - \alpha.$$

Aufgrund der stochastischen Konvergenz der Zufallsvariablen $\hat{a}_{ML}$ gegen ∞ für $a \rightarrow \infty$ (s. Kapitel 2.6) kann damit für alle Funktionen c , die

$$\lim_{\hat{a} \rightarrow \infty} c(\hat{a}) = c_\infty$$

erfüllen, der Grenzwert

$$\lim_{a \rightarrow \infty} P_{a, \sigma^2} \left\{ \sigma^2 \in C(\underline{X}) \right\} = 1 - \alpha \quad (4.58)$$

gebildet werden. Damit ist auch die in der Interpolationsvorschrift angegebene Wahl der Konstanten $\gamma_{104} = c_\infty$ offenkundig.

4.10. Einige Beispiele

In diesem Abschnitt sollen Abbildungen der Funktionen c für die Stichprobenumfänge $n = 10, 20, 50, 250$ und Konfidenzniveaus $1 - \alpha = 0,90; 0,95; 0,99$ angegeben werden. Die zur Bestimmung dieser Konfidenzintervalle benötigte Zahl von Iterationsschritten sowie die erzielte „Genauigkeit“ kann aus der Tabelle 4.3 (Abschnitt 4.9) entnommen werden, wobei jeweils so viele Schritte durchgeführt wurden, daß die betragsmäßig größte Abweichung

$$\max_{a \in A} |\tilde{\alpha}^{(i)}(a) - \alpha|$$

minimiert wird.

Der Ansatz

Da es — wie in Abb. 4.17 zu erkennen ist — nicht möglich ist die Originalfunktionen c für mehrere Stichprobenumfänge in angemessener Form in einer Zeichnung unterzubringen, wird eine Verfeinerung des Interpolationsansatzes vorgenommen, die eine gute Darstellung der Funktionen ermöglicht. Der aus Abschnitt 4.9 übernommene Ansatz lautet

$$c(\hat{a}) = \gamma e^{\frac{\hat{a}}{4}}. \quad (4.59)$$

In Abb. 4.18 wurde die „Güte“ dieses Ansatzes illustriert. Da dieser Ansatz sicherlich nicht exakt ist, soll er zu

$$c(\hat{a}) = c_{n,\alpha} e^{\frac{\delta(\hat{a})}{4}} \quad (4.60)$$

erweitert werden, in der durch die Abb. 4.18 unterstützten Erwartung, daß die Funktion δ nahezu konstant ist. Die Konstante $c_{n,\alpha}$ stellt dabei die im letzten Abschnitt als c_∞ bezeichnete Konstante

$$c_{n,\alpha} \stackrel{\text{def}}{=} \frac{n}{\mathcal{G}_{\frac{(n-1)}{2}, 2(n-1)}^{-1}(\alpha)}$$

dar, welche, wie dort gezeigt, ein geeigneter Kandidat für γ ist. Wie die folgenden Abbildungen zeigen werden, ergibt sich eine weitere Möglichkeit, den Ansatz zu verfeinern. Die sich aus der Umformung

$$\hat{a} \left(\frac{\ln c(\hat{a})}{\ln c_{n,\alpha}} - 1 \right) = \frac{\delta(\hat{a})}{\ln c_{n,\alpha}}$$

ergebende Größe $\frac{\delta(\hat{a})}{\ln c_{n,\alpha}}$ hängt nur noch in äußerst geringem Maße vom Stichprobenumfang n und vom Konfidenzniveau ab. Dieses bedeutet, daß im Ansatz

$$c(\hat{a}) = (c_{n,\alpha})^{1 + \frac{\delta^*(\hat{a})}{4}}, \quad (4.61)$$

der mit $\delta^*(\hat{a}) \stackrel{\text{def}}{=} \frac{\delta(\hat{a})}{\ln c_{n,\alpha}}$ identisch zum Ansatz (4.60) ist, die Funktion δ^* für alle Werte von $\hat{a}$, n und α nahezu konstant ist. Die Darstellung (4.61) vermittelt zudem einen guten Eindruck von der Gestalt der Funktion $c(\hat{a})$. Der Faktor $(c_{n,\alpha})^{\frac{\delta^*(\hat{a})}{\hat{a}}}$ ist sozusagen ein „Korrekturfaktor“ für $\hat{a} < \infty$ (die im letzten Abschnitt beschriebene Interpolationsvorschrift besagt, daß $\frac{\delta^*(\hat{a})}{\hat{a}}$ für $\hat{a} \rightarrow \infty$ gegen 0 konvergiert). Aus der Beobachtung, daß δ^* nur relativ geringfügig von $\hat{a}$ abhängt, ist eine gute Beurteilungsmöglichkeit des Ausmaßes der Korrektur möglich. Zudem könnte versucht werden, theoretische Aussagen für δ^* herzuleiten. Da die Funktion δ^* nur geringfügig von n und α abhängt, ist es überdies denkbar, durch eine numerische Anpassung der Funktion δ^* eine einfache funktionale Gestalt für c zu bestimmen, welche dennoch eine hohe Güte für einen sehr großen Bereich von Werten von n und α besitzt.

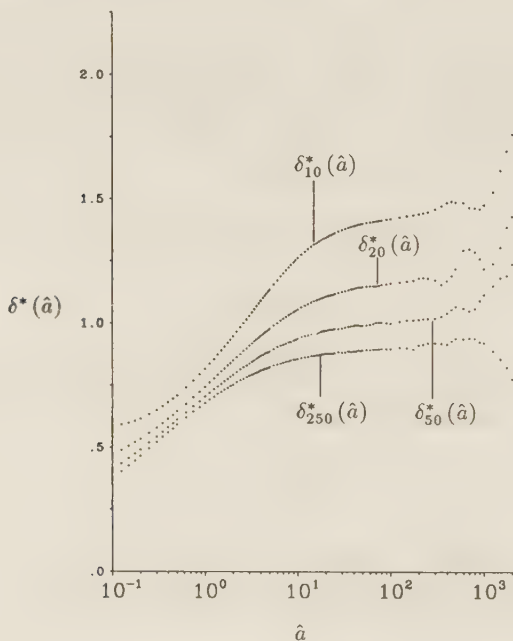


Abb. 4.20 δ^* für $1 - \alpha = 0,95$ und $n = 10, 20, 50, 250$

Die Werte der numerischen Berechnungen der Funktion δ^* streuen für große Werte von $\hat{a}$ relativ stark, wie in Abb. 4.20 zu erkennen ist. Dieses ist allerdings unproblematisch, die Schwankungsbreite der Funktion δ^* hat, wie durch Einsetzen in (4.61) sofort ersichtlich wird, auf den Funktionswert von c für große Werte von $\hat{a}$ keinen nennenswerten Einfluß — und damit auch nicht auf den Funktionswert

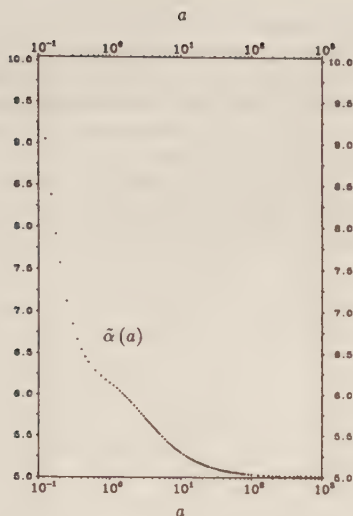


Abb. 4.21 Die Funktion $\tilde{\alpha}$ für $n = 10$ unter Verwendung von δ_{250}^*

von $\tilde{\alpha}$ für große Werte von a . Umgekehrt formuliert heißt dies, daß schon kleinste Schwankungen des Funktionswertes von c für große Werte von $\hat{a}$ relativ große Schwankungen des Funktionswertes von δ^* zur Folge haben. Werden die zur Bestimmung der Funktionen δ^* in den Abbildungen 4.20, 4.23 und 4.24 verwendeten, auf sechs Dezimalstellen genauen Werte von $c_{n,\alpha}$ auf z.B. vier Dezimalstellen gerundet, so ergeben sich im Bereich mittlerer und großer Werte von $\hat{a}$ deutlich andere Funktionsverläufe von δ^* , die Funktionen schneiden sich teilweise schon im Bereich $\hat{a} = 100$. Es ist also absolut ausreichend, für große Werte von $\hat{a}$ die Größenordnung von δ^* bestimmen zu können, um eine hohe Genauigkeit des Funktionswertes von c zu erhalten (in Abschnitt 4.9 wurde gezeigt, daß bereits aus $\lim_{\hat{a} \rightarrow \infty} \frac{\delta^*(\hat{a})}{\hat{a}} = 0$ der Grenzwert $\lim_{a \rightarrow \infty} P_{a,\sigma^2} \{ \sigma^2 \in C(\underline{X}) \} = 1 - \alpha$ folgt).

Der Einfluß des Stichprobenumfangs

Mit Hilfe der in Abb. 4.20 gezeichneten Kurven ist es bereits möglich, c auch für andere Stichprobenumfänge so zu bestimmen, daß die Funktion $\tilde{\alpha}$ nahe an der Konstanten $\alpha = 0,05$ liegt, es reicht also, die Funktion δ^* für sehr wenige Stichprobenumfänge zu berechnen. Einfache Interpolationsalgorithmen ergeben bereits ein im Sinne der Einhaltung des Konfidenzniveaus sehr gutes Konfidenzintervall. Diese Eigenschaft, die einen hohen praktischen Wert hat, soll an zwei Extrembeispielen verdeutlicht werden. In Abbildung 4.21 ist die Funktion $\tilde{\alpha}$ dargestellt, die sich für $n = 10$ ergibt, wenn anstelle der zu $n = 10$ gehörenden Funktion δ_{10}^* die für $n = 250$ berechnete Funktion δ_{250}^* in (4.61) eingesetzt wird.

Es zeigt sich, daß $\tilde{\alpha}$ im gesamten Bereich von A im Bereich $(0,05; 0,10)$ liegt, für $a > 0,25$ beträgt der Wert von $\tilde{\alpha}$ bereits weniger als 0,07, und für $a > 1,5$ weniger als 0,06. An der Stelle $a = 300$ wird $\tilde{\alpha}(a) = 0,0501$ erreicht, ehe die Funktion wieder bis auf 0,0507 ansteigt. Dieser wohl ungünstigste Ansatz, bei dem immerhin ein um den Faktor 25 zu großer Stichprobenumfang zur Berechnung der Funktion δ^* angesetzt wurde, zeigt, daß die Funktion δ^* im Sinne der Einhaltung des Konfidenzniveaus nur geringfügig von n abhängt. In Abb. 4.22 wird zur weiteren Illustration noch diejenige Funktion $\tilde{\alpha}$ angegeben, die sich aus der Verwendung der zu $n = 50$ gehörenden Funktion δ_{50}^* im Falle $n = 250$ ergibt.

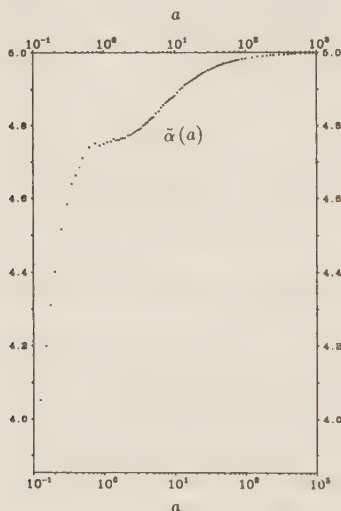


Abb. 4.22 Die Funktion $\tilde{\alpha}$ für $n = 250$ unter Verwendung von δ_{50}^*

Der Einfluß des Konfidenzniveaus

In den Abb. 4.23 und 4.24 sind die Funktionen δ^* für die Stichprobenumfänge $n = 10, 20, 50, 250$ und die Konfidenzniveaus 0,90 sowie 0,99 gezeichnet. Der geringe Einfluß des Konfidenzniveaus auf die Funktion δ^* soll ebenfalls durch zwei extreme Beispiele illustriert werden. In Abb. 4.25 wird für $n = 50$ die Funktion $\tilde{\alpha}$ gezeichnet, die sich ergibt, wenn die zum Niveau 0,99 konstruierte Funktion $\delta_{0,99}^*$ über (4.61) zur Konstruktion des Konfidenzintervalls zum Niveau $1 - \alpha = 0,90$ herangezogen wird, während in Abb. 4.26 das Resultat der umgekehrten Vorgehensweise dargestellt wird. Deutlich zu erkennen ist auch hier, daß die Annäherung an die Konstante α für nicht zu kleine Werte des Parameters a selbst in diesen extremen Beispielen recht gut ist.

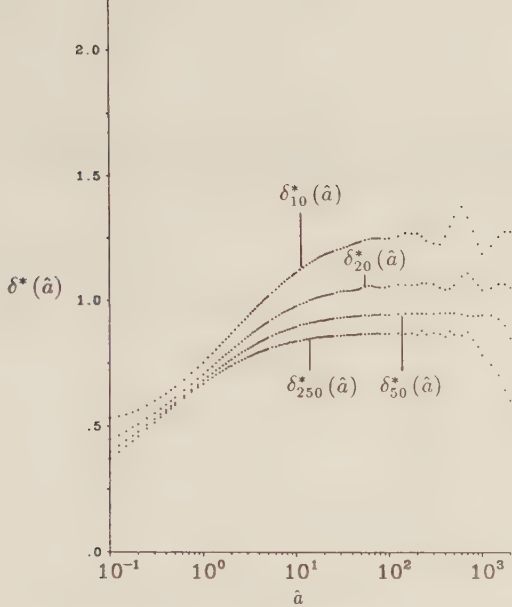


Abb. 4.23 δ^* für $1 - \alpha = 0,90$ und $n = 10, 20, 50, 250$

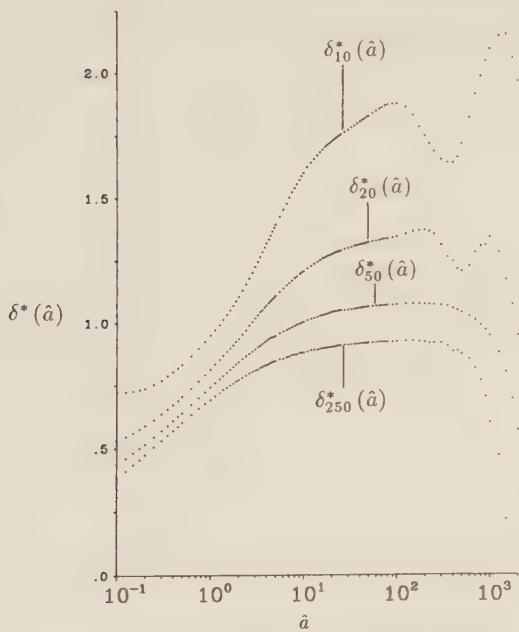


Abb. 4.24 δ^* für $1 - \alpha = 0,99$ und $n = 10, 20, 50, 250$

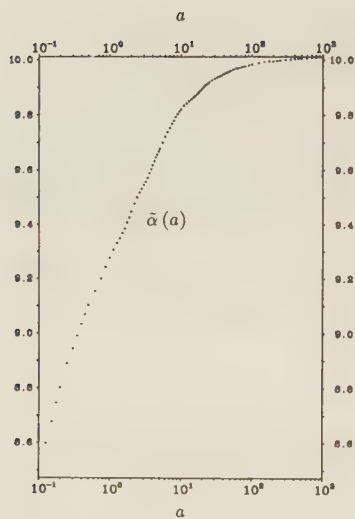


Abb. 4.25 Die Funktion $\tilde{\alpha}$ für $1 - \alpha = 0,90$ unter Verwendung von $\delta_{0,99}^*$

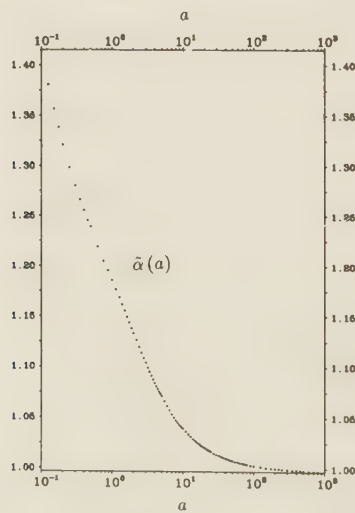


Abb. 4.26 Die Funktion $\tilde{\alpha}$ für $1 - \alpha = 0,99$ unter Verwendung von $\delta_{0,90}^*$

4.11. Überlegungen zu Verallgemeinerungsmöglichkeiten

Im folgenden sollen kurz einige Überlegungen zur Übertragbarkeit des in diesem Kapitel vorgestellten Verfahrens auf andere Intervallschätzprobleme dargestellt werden.

Exponentialfamilie

Die zugrundeliegende Verteilungsfamilie soll dabei stets eine zweiparametrische Exponentialfamilie mit der Dichtefunktion

$$f_{\theta, \delta}(x) = g(\theta, \delta) h(x) e^{\theta T(x) + \delta U(x)}$$

sein. Für θ und δ (bei unbekanntem zweiten Parameter) existieren gleichmäßig beste einseitige, bzw. in der Klasse der unverfälschten, gleichmäßig beste zweiseitige Tests (s. Lehmann (1986, S. 145ff)), aus denen sich aufgrund der Dualität Testfamilie-Konfidenzintervall Konfidenzintervalle mit den entsprechenden Optimalitätseigenschaften ergeben. Konfidenzintervalle für θ werden mit Hilfe der bedingten Verteilung von T gegeben $U = u$, die nicht von δ abhängt, gebildet, Konfidenzintervalle für δ über die bedingte Verteilung von U gegeben $T = t$, die nicht von θ abhängt. Damit bleibt in diesen Fällen „nur“ noch das Problem, Quantile der jeweiligen bedingten Verteilung zu berechnen. Soll allerdings ein Konfidenzintervall für eine Funktion von θ und δ gebildet werden, so existieren keine allgemeinen theoretischen Aussagen mehr. Im vorliegenden Gammaverteilungsmodell gilt $\theta = a$ und $\delta = \frac{1}{b}$, so daß Konfidenzintervalle für a und $\frac{1}{b}$ (und damit natürlich auch für b) existieren, die die oben erwähnten Optimalitätskriterien erfüllen. Eine Reihe von Arbeiten beschäftigt sich mit der Bildung solcher Tests bzw. Konfidenzintervalle: Shorack (1972), Bar-Lev, Reiser (1982), Engelhardt, Bain (1977, 1978), Lentner, Buehler (1963), Linhardt (1965) und Moran (1951).

Für die Varianz $ab^2 = \frac{\theta}{\delta^2}$ oder den Erwartungswert $ab = \frac{\theta}{\delta}$ der Gammaverteilung liefert die Theorie jedoch keine Aussagen. In der Familie der Normalverteilungen mit Erwartungswert μ und Varianz σ^2 gilt $\theta = \frac{1}{\sigma^2}$ und $\delta = \frac{\mu}{\sigma^2}$, so daß das bekannte Konfidenzintervall für σ^2 durch die obige theoretische Verfahrensweise konstruierbar ist, Aussagen über die Konstruktion eines Konfidenzintervalls für μ liefert der oben erwähnte Sachverhalt schon nicht mehr. Die Reihe der Beispiele läßt sich noch fortsetzen, genannt werden sollen an dieser Stelle nur noch die Fragestellung der Bestimmung eines Konfidenzintervalls für den Erwartungswert oder die Varianz der Betaverteilung.

Zur Gammaverteilungssituation

Im Falle der Bestimmung eines Konfidenzintervalls für θ und δ wird das Problem, daß ein zweiter — unbekannter — Parameter vorliegt, durch die Anwendung des oben erwähnten bedingten Tests bzw. Konfidenzintervalls gelöst. Bei der in diesem Kapitel behandelten Fragestellung schuf aber gerade die Tatsache, daß ein Nuisance-Parameter — der Parameter a — existiert, erhebliche Schwierigkeiten. Das beschriebene Verfahren kann in völliger Analogie auch auf die Problemstellung der Konstruktion eines einseitigen Konfidenzintervalls für andere Parameter der

Gammaverteilung — etwa den Erwartungswert — übertragen werden. Auch hier werden natürlich die für die Klasse der Gammaverteilungen suffizienten Variablen $\bar{X}$ und Q verwendet, die in den Abschnitten 4.1–4.4 sowie 4.7–4.9 gemachten Aussagen lassen sich direkt übertragen, wenn anstelle der Zufallsvariablen $\frac{X}{\sigma}$ und deren Verteilungsfunktion G_a , die Zufallsvariable $\frac{X}{E(\bar{X})}$, die die Verteilungsfunktion $G_{a, \frac{1}{4}}$ besitzt, verwendet wird. Da die Verteilung von $\hat{a}_{ML}$ weder von der Varianz noch vom Erwartungswert, sondern nur von a abhängt, bleibt auch hier das Problem der „Elimination“ des Parameters a . Selbst das numerische Verhalten des Verfahrens, welches, wie in Abschnitt 4.8 erläutert wurde, vom Einfluß des Parameters a auf die beteiligten Verteilungen bestimmt wird, dürfte vergleichbar sein.

Zur Übertragung auf andere Verteilungsfamilien

Das Verfahren kann prinzipiell auch auf andere Exponentialfamilien übertragen werden. Dabei bleiben aber im Einzelfall einige Fragen zu klären. Diese sollen anhand der Übertragung auf die Fragestellung der Bestimmung eines einseitigen Konfidenzintervalls für den Erwartungswert μ der Normalverteilung verdeutlicht werden. Natürlich ist es nicht sehr sinnvoll, ein numerisches Verfahren auf eine bereits theoretisch gelöste Problemstellung anzuwenden, doch es soll an dieser Stelle nur die prinzipielle Übertragbarkeit verdeutlicht werden, wobei es sich zeigt, daß das Verfahren — von numerischen Abweichungen abgesehen — ohne daß spezielle theoretische Aussagen über das Student-Konfidenzintervall in das Verfahren eingehen, als Ergebnis exakt dieses konstruiert hätte und zwar bereits durch die Startfunktion. Die Übertragung soll im folgenden nur skizziert werden, es werden also nur knappe Hinweise gegeben, die die Analogien zwischen den Konfidenzintervallen verdeutlichen sollen.

Aufgrund einer Stichprobe $\underline{X} = (X_1, \dots, X_n)$ von unabhängigen und identisch normalverteilten Zufallsvariablen soll bei unbekannter Varianz σ^2 ein einseitiges Konfidenzintervall

$$C(\underline{X}) = \{\mu : \mu \leq c(\underline{X})\}$$

für den Erwartungswert μ zum Konfidenzniveau $1 - \alpha$ gebildet werden. Ebenso wie in der Gammaverteilungssituation lassen sich zwei unabhängige und für die Familie der Normalverteilungen suffiziente Zufallsvariable angeben: $\bar{X}$ und die Stichprobenvarianz S^2 . Die Verteilung von S^2 hängt dabei wie die Verteilung von Q im Gammaverteilungsmodell nur vom Nuisance-Parameter ab. Der zu schätzende Parameter μ läßt sich durch den Übergang zu der Transformation $(\bar{X} - \mu, S^2)$ ebenso wie der in der Gammaverteilungssituation zu schätzende Parameter σ^2 durch $\left(\frac{\bar{X}}{\sigma}, Q\right)$ im Rahmen einer Pivotgrößenbildung behandeln, so daß die Übertragung der Annahmen aus Abschnitt 4.2 die Form

$$c(\underline{X}) = c(S^2) + \bar{X}$$

liefert. Der in Abschnitt 4.7 eingeführten Familie $\{h_a^* : a > 0\}$ entspricht hier eine Familie $\{h_{\sigma^2}^* : \sigma^2 > 0\}$, die Familie der ξ -Funktionen läßt sich ebenfalls entsprechend bilden. Die in Abschnitt 4.8 getroffenen Pivotannahmen, die wie erläutert

— falls sie erfüllt wären — bedeuteten, daß die Startfunktion $c^{(1)}$ bereits ein ähnliches Konfidenzintervall definierte, sind — auf die Normalverteilungssituation übertragen — erfüllt, da es gelingt, den Nuisance-Parameter σ^2 durch den Übergang zum Tupel $\left(\frac{\bar{X}-\mu}{\sigma}, \frac{S^2}{\sigma^2}\right)$ zu „eliminieren“. Dieses bedeutet, daß die zur Regel (4.45) äquivalente Konstruktionsbedingung für die Startfunktion $c^{(1)}$ zum Erfolg führt, das Verfahren konstruiert in der Normalverteilungssituation — wie sich analog zu den Ausführungen des Abschnitts 4.8 zeigen läßt — eine Startfunktion $c^{(1)}$ der Gestalt

$$c(S^2) = t\sqrt{S^2},$$

wobei die Konstante t numerisch so bestimmt wird, daß die Integralbedingung

$$\alpha = \int_0^1 \Phi\left(t\sqrt{\frac{n}{n-1}G_{\left(\frac{n-1}{2}, 2(n-1)\right)}^{-1}}(u)\right) du$$

erfüllt ist (Φ : Verteilungsfunktion der Standardnormalverteilung). Die Lösung dieser Gleichung läßt sich auch theoretisch angeben:

$$t = \frac{t_{n-1}(1-\alpha)}{\sqrt{n}},$$

wobei $t_{n-1}(1-\alpha)$ das $(1-\alpha)$ -Quantil der t -Verteilung mit $n-1$ Freiheitsgraden bezeichnet. Damit würde bei einer Übertragung des Verfahrens auf die gerade beschriebene Normalverteilungssituation bereits durch die Wahl der Startfunktion $c^{(1)}$ das Konfidenzintervall

$$C(\underline{X}) = \left\{ \mu : \mu \leq \bar{X} + t_{n-1}(1-\alpha)\sqrt{\frac{S^2}{n}} \right\},$$

also das einseitige t -Konfidenzintervall, konstruiert werden.

Zweiseitige Konfidenzintervalle

Bei der Übertragung des Verfahrens auf zweiseitige Fragestellungen ergibt sich ein zusätzliches Problem. In der Situation der Bestimmung des t -Konfidenzintervalls ist es möglich durch die Bestimmung je eines einseitigen zu den Konfidenzniveaus $1 - \frac{\alpha}{2}$ und $\frac{\alpha}{2}$ konstruierten Konfidenzintervalls zum Ziel zu kommen. Das erste Konfidenzintervall liefert die Obergrenze, das zweite die Untergrenze des gesuchten zweiseitigen Intervalls. Dieses Vorgehen ist auch auf andere Situationen übertragbar, doch ist das so konstruierte Konfidenzintervall im allgemeinen verfälscht. In der Gammaverteilungssituation wäre es nötig, zwei einseitige Konfidenzintervalle zu konstruieren, eines zum Niveau $1 - \alpha_1(a)$ und eines zum Niveau $\alpha_2(a)$ mit $\alpha = \alpha_1(a) + \alpha_2(a)$. Natürlich ist es möglich den Prozeß der iterativer Anpassung der Funktion $\tilde{\alpha}$ an die Konstante α auch auf die Anpassung an die Funktionen $\alpha_i(a)$, $i = 1, 2$ zu übertragen, doch es ist eben unbekannt, wie diese Funktionen aussehen (für große Werte von a wird wohl $\alpha_i(a) \approx \frac{\alpha}{2}$, $i = 1, 2$ gelten). Es ergibt sich also — falls ein unverfälschtes Konfidenzintervall gebildet werden soll — die zusätzliche Problematik der Konstruktion dieser beiden Funktionen.

5. Ein Verfahren zur Abschätzung der Überdeckungswahrscheinlichkeiten

Einige Vorbemerkungen

In diesem Kapitel soll die in Kapitel 4.1 angesprochene Abschätzung der Funktion $\tilde{\alpha}$ für Werte des Parameters a , die außerhalb des Stützstellenbereichs A liegen, angegeben werden. Dazu werden möglichst gute Abschätzungen der Verteilungsfunktionen $\mathcal{G}_{an}$ und $\mathcal{A}_{a,n}$ benötigt, wobei die Funktionen wiederum als Funktionen des Parameters a aufgefaßt werden. Die Fragestellung besteht in der Angabe je einer für alle $a \in (a_1, a_2)$ gültigen Abschätzung des Integrals

$$\int_0^1 h_a(u) du \quad (5.1)$$

nach unten und oben.

Bei der Konstruktion einer geeigneten Abschätzung erweist es sich als wichtig, die Zufallsvariablen $\bar{X}$ und $\hat{a}_{ML}$ so zu transformieren, daß die Verteilungsfunktionen der transformierten Variablen nur noch „geringfügig“ vom Parameter a abhängen. Aus den in den Abschnitten 5.1 und 5.2 hergeleiteten relativ scharfen Abschätzungen der Verteilungen der transformierten Variablen können dann durch Rücktransformation entsprechend gute Abschätzungen der Verteilungsfunktionen von $\bar{X}$ und $\hat{a}_{ML}$ bestimmt werden. Im dritten Abschnitt wird schließlich die Abschätzung des Zielintegrals beschrieben. Analog zur dort beschriebenen Vorgangsweise können auch Abschätzungen aus den nachfolgend angegebenen Monotonieeigenschaften der nichttransformierten Zufallsvariablen $\bar{X}$ und $\hat{a}_{ML}$

$$\begin{aligned} \forall \hat{a} > 0 \quad \forall a > 0: \quad \frac{\partial}{\partial a} P_a \{ \hat{a}_{ML} \leq \hat{a} \} &= \frac{\partial}{\partial a} \mathcal{A}_{a,n}(\hat{a}) < 0 \\ \forall x > 0 \quad \forall a, b > 0: \quad \frac{\partial}{\partial a} P_{a,\sigma^2} \{ \bar{X} \leq x \} &= \frac{\partial}{\partial a} \mathcal{G}_{an} \left(x \sqrt{\frac{n}{ab^2}} \right) < 0 \end{aligned}$$

gewonnen werden, doch sind diese Abschätzungen viel zu grob, um einigermaßen befriedigende Ergebnisse zu liefern. Der Beweis der oberen Beziehung wurde in Kapitel 2.8 erbracht, auf den einfachen Beweis der zweiten Aussage soll hier verzichtet werden. Da die Fehler, der sich aus diesen beiden Monotonieeigenschaften ergebenden Abschätzungen, von der Ordnung $o\left(\left(\frac{a_2}{a_1}\right)^n\right)$ sind, soll ohnehin auf eine weitere Diskussion verzichtet werden.

Anstelle der Variablen $\bar{X}$ und $\hat{a}_{ML}$ werden die Verteilungen der bijektiven Transformationen $\frac{\bar{X}}{E(\bar{X})}$ und Q^a betrachtet. Die Zufallsvariable Q^a ihrerseits ist eine monotone Transformation der Zufallsvariablen $-2an \ln Q = -2n \ln(Q^a)$, deren Verteilung in Kapitel 2.5 ausführlich behandelt wurde. Insbesondere zeigte

sich, daß sich beim Grenzübergang $a \rightarrow 0$ bzw. $a \rightarrow \infty$ χ^2 -Verteilungen mit $2(n-1)$ bzw. $n-1$ Freiheitsgraden ergeben, was für nicht zu kleine n eine recht vielversprechende Eigenschaft ist.

Im folgenden wird stets $\delta \stackrel{\text{def}}{=} \frac{a_2}{a_1}$ verwendet, an die Stelle der Formulierung $\forall a \in (a_1, a_2)$ tritt unter Einführung von $\epsilon \stackrel{\text{def}}{=} \frac{a}{a_1}$ die äquivalente Sprechweise $\forall \epsilon \in (1, \delta)$.

5.1. Die Abschätzung der Verteilungsfunktion von $\bar{X}$

Die Abschätzung der Verteilungsfunktion von $\bar{X}$ kann mit Hilfe von Abschätzungen der Dichtefunktion der Zufallsvariablen $\frac{\bar{X}}{E(\bar{X})}$ vorgenommen werden.

Sei X eine $G(a, b)$ -verteilte Zufallsvariable. Aus den in Kapitel 2.1 angegebenen Zusammenhängen folgt, daß die Zufallsvariable $Z \stackrel{\text{def}}{=} \frac{X}{E(X)}$ ebenfalls gammaverteilt ist mit Gestaltparameter a und Skalierungsparameter $\frac{1}{a} = \frac{b}{ab}$. Die Varianz von Z beträgt $\frac{1}{a} = \frac{ab^2}{(ab)^2}$, die Dichtefunktion $f_a(z)$ lautet für $z > 0$

$$\forall a > 0: \quad f_a(z) = \frac{a^a}{\Gamma(a)} z^{a-1} e^{-az}. \quad (5.2)$$

Die Verteilungsfunktion von Z soll mit F_a bezeichnet werden. Der Übersichtlichkeit halber wird im folgenden der Index 1 der oben eingeführten Konstanten a_1 unterdrückt, es werden also für alle $\epsilon \in (1, \delta)$ gültige Abschätzungen der Dichtefunktion $f_{a\epsilon}$ gesucht. Differenzieren des für $z > 0$ definierten Dichtequotienten

$$q_z(\epsilon) \stackrel{\text{def}}{=} \frac{f_{a\epsilon}(z)}{f_a(z)} \quad (5.3)$$

nach z liefert

$$\frac{\partial}{\partial z} q_z(\epsilon) = \left(\frac{1}{z} - 1 \right) q_z(\epsilon). \quad (5.4)$$

Der Dichtequotient ist also für $z < 1$ isoton und für $z > 1$ antiton. Mit der Normierungsbedingung

$$\int_0^\infty f_{a\epsilon}(z) dz = 1 = \int_0^\infty f_a(z) dz$$

folgt für $\epsilon \in (1, \delta)$ die Existenz von $z_1(\epsilon) < 1$ und $z_2(\epsilon) > 1$ mit

$$q_z(\epsilon) > 1 \quad \Longleftrightarrow \quad z_1(\epsilon) < z < z_2(\epsilon). \quad (5.5)$$

Ableiten des Dichtequotienten nach ϵ führt zu

$$\frac{\partial}{\partial \epsilon} q_z(\epsilon) = a q_z(\epsilon) (1 + \ln z - z - \rho(a\epsilon)), \quad (5.6)$$

wobei ρ die in Kapitel 2.4 diskutierten Eigenschaften besitzt, also negativ und streng isoton ist mit

$$\lim_{x \rightarrow 0} \rho(x) = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \rho(x) = 0,$$

während $1 + \ln z - z$ an der Stelle $z = 1$ das Maximum 0 besitzt und für $z < 1$ isoton sowie für $z > 1$ antiton ist.

Mit $q_z(1) = 1$ ergeben sich daraus für $1 < \epsilon < \delta$ die folgenden Fallunterscheidungen, die in der nachfolgenden Abb. 5.1 noch einmal verdeutlicht werden sollen:

$$\begin{aligned} \text{I} \quad 1 + \ln z - z &\geq \rho(a\delta) : & \frac{\partial}{\partial \epsilon} q_z(\epsilon) &\geq 0 \implies f_a(z) \leq f_{a\epsilon}(z) \leq f_{a\delta}(z) \\ \text{II} \quad \rho(a\delta) &> 1 + \ln z - z > \rho(a) : & q_z(\epsilon) &\text{ besitzt ein Maximum für ein } \epsilon_0 \in (1, \delta) \\ \text{III} \quad \rho(a) &\geq 1 + \ln z - z : & \frac{\partial}{\partial \epsilon} q_z(\epsilon) &\leq 0 \implies f_{a\delta}(z) \leq f_{a\epsilon}(z) \leq f_a(z) \end{aligned} \quad (5.7)$$

Der Fall II muß noch weiter unterschieden werden:

$$\begin{aligned} \text{II a} \quad q_z(\delta) &\geq 1 : & f_a(z) &\leq f_{a\epsilon}(z) \\ \text{II b} \quad q_z(\delta) &< 1 : & f_{a\delta}(z) &\leq f_{a\epsilon}(z) \end{aligned} \quad (5.8)$$

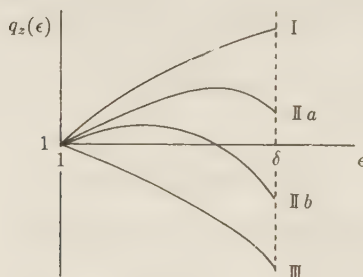
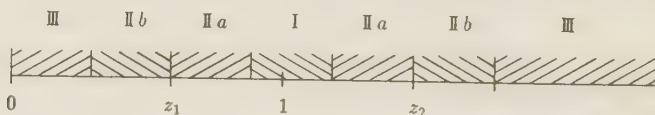


Abb. 5.1 Die Funktion $q_z(\epsilon)$

Die Anordnung dieser Bereiche läßt sich am Zahlenstrahl wie folgt darstellen, wobei $z_1 \stackrel{\text{def}}{=} z_1(\delta)$ und $z_2 \stackrel{\text{def}}{=} z_2(\delta)$ (s. 5.5) gesetzt wird:



Es fehlt noch eine Abschätzung des Dichtequotienten nach oben für den Bereich II.

Bevor dieses Problem betrachtet wird, können aus den für alle $z > 0$ bereits angegebenen unteren Schranken der Dichtefunktion bereits Abschätzungen der Verteilungsfunktion $F_{a\epsilon}(z)$ bestimmt werden. Durch stückweise Integration und unter Verwendung von $F_{a\epsilon}(z) = \int_0^z f_{a\epsilon}(y) dy$ ergeben sich Abschätzungen nach unten und unter Verwendung von $F_{a\epsilon}(z) = 1 - \int_z^\infty f_{a\epsilon}(y) dy$ Abschätzungen nach oben:

Bestimme $z_1 < 1$ und $z_2 > 1$, so daß gilt:

$$f_{a\delta}(z_i) = f_a(z_i), \quad i = 1, 2.$$

1. Abschätzung nach unten:

$$0 < z \leq z_1 : F_{a\epsilon}(z) \geq F_{a\delta}(z)$$

$$z_1 < z \leq z_2 : F_{a\epsilon}(z) \geq F_{a\delta}(z_1) - F_a(z_1) + F_a(z)$$

$$z_2 < z : F_{a\epsilon}(z) \geq F_{a\delta}(z_1) - F_a(z_1) + F_a(z_2) - F_{a\delta}(z_2) + F_{a\delta}(z) \quad (5.9a)$$

1. Abschätzung nach oben:

$$0 < z \leq z_1 : F_{a\epsilon}(z) \leq F_{a\delta}(z_2) - F_a(z_2) + F_a(z_1) - F_{a\delta}(z_1) + F_a(z)$$

$$z_1 < z \leq z_2 : F_{a\epsilon}(z) \leq F_{a\delta}(z_2) - F_a(z_2) + F_a(z)$$

$$z_2 < z : F_{a\epsilon}(z) \leq F_{a\delta}(z) \quad (5.9b)$$

Die Abschätzung nach unten wird insbesondere im Bereich großer, die Abschätzung nach oben im Bereich kleiner Werte von z relativ schlecht sein, so daß aus den oberen Schranken der Dichtefunktion Abschätzungen der Verteilungsfunktion angegeben werden sollen, für die genau die umgekehrte Aussage zutrifft.

An den beiden Grenzen der Bereiche III und II b, also für $1 + \ln z - z = \rho(a)$ gilt

$$\sup_{1 < \epsilon < \delta} q_z(\epsilon) = q_z(1) = 1,$$

an den Grenzen der Bereiche II a und I, also für $1 + \ln z - z = \rho(a\delta)$

$$q^* \stackrel{\text{def}}{=} \sup_{1 < \epsilon < \delta} q_z(\epsilon) = q_z(\delta) = \frac{\Gamma(a)}{\Gamma(a\delta)} \delta^{a\delta} a^{a(\delta-1)} e^{a(\delta-1)(\rho(a\delta)-1)}. \quad (5.10)$$

Aus der Monotonieeigenschaft (5.4) folgt zusammen mit $q^* > 1$, daß in beiden Bereichen II der Dichtequotient $q_z(\epsilon)$ stets durch q^* nach oben abgeschätzt werden kann:

$$q_z(\epsilon_0) = \sup_{1 < \epsilon < \delta} q_z(\epsilon) \leq q^*.$$

Die folgende Abschätzung (5.11) ist nicht so schwach wie es auf den ersten Blick erscheinen mag, denn die Dichtefunktion $f_{a\epsilon}(z)$ wird nicht direkt, sondern indirekt über den Dichtequotienten $q_z(\epsilon)$ abgeschätzt:

$$\text{II} \quad \rho(a\delta) > 1 + \ln z - z > \rho(a) : \quad f_{a\epsilon}(z) \leq q_z(\epsilon_0) f_a(z) \leq q^* f_a(z) \quad (5.11)$$

Aus (5.7) und (5.11) ergeben sich damit wiederum durch stückweise Integration die folgenden für $\epsilon \in (1, \delta)$ gültigen Abschätzungen:

Bestimme $z_3 < 1$ und $z_6 > 1$, so daß

$$1 + \ln z_i - z_i = \rho(a), \quad i = 3, 6$$

gilt, sowie $z_4 < 1$ und $z_5 > 1$, so daß

$$1 + \ln z_i - z_i = \rho(a\delta), \quad i = 4, 5$$

gilt. Setze

$$q^* \stackrel{\text{def}}{=} \frac{\Gamma(a)}{\Gamma(a\delta)} \delta^{a\delta} a^{a(\delta-1)} e^{a(\delta-1)(\rho(a\delta)-1)}.$$

2. Abschätzung nach unten:

$$0 < z \leq z_3 : \quad F_{a\epsilon}(z) \geq (1 - q^*)F_a(z_6) + q^*F_a(z_5) - F_{a\delta}(z_5) + F_{a\delta}(z_4) \\ - q^*F_a(z_4) + (q^* - 1)F_a(z_3) - F_a(z)$$

$$z_3 < z \leq z_4 : \quad F_{a\epsilon}(z) \geq (1 - q^*)F_a(z_6) + q^*F_a(z_5) - F_{a\delta}(z_5) + F_{a\delta}(z_4) \\ - q^*F_a(z_4) + q^*F_a(z)$$

$$z_4 < z \leq z_5 : \quad F_{a\epsilon}(z) \geq (1 - q^*)F_a(z_6) + q^*F_a(z_5) - F_{a\delta}(z_5) + F_{a\delta}(z)$$

$$z_5 < z \leq z_6 : \quad F_{a\epsilon}(z) \geq (1 - q^*)F_a(z_6) + q^*F_a(z)$$

$$z_6 < z : \quad F_{a\epsilon}(z) \geq F_a(z)$$

(5.12a)

2. Abschätzung nach oben:

$$0 < z \leq z_3 : \quad F_{a\epsilon}(z) \leq F_a(z)$$

$$z_3 < z \leq z_4 : \quad F_{a\epsilon}(z) \leq (1 - q^*)F_a(z_3) + q^*F_a(z)$$

$$z_4 < z \leq z_5 : \quad F_{a\epsilon}(z) \leq (1 - q^*)F_a(z_3) + q^*F_a(z_4) - F_{a\delta}(z_4) + F_{a\delta}(z)$$

$$z_5 < z \leq z_6 : \quad F_{a\epsilon}(z) \leq (1 - q^*)F_a(z_3) + q^*F_a(z_4) - F_{a\delta}(z_4) + F_{a\delta}(z_5) \\ - q^*F_a(z_5) + q^*F_a(z)$$

$$z_6 < z : \quad F_{a\epsilon}(z) \leq (1 - q^*)F_a(z_3) + q^*F_a(z_4) - F_{a\delta}(z_4) + F_{a\delta}(z_5) \\ - q^*F_a(z_5) + (q^* - 1)F_a(z_6) + F_a(z)$$

(5.12b)

Im Bereich großer Werte von z wird die Abschätzung nach unten, im Bereich kleiner Werte wird die Abschätzung nach oben relativ gut sein. Für ein gegebenes

z kann aus (5.9) und (5.12) die jeweils bessere Abschätzung gewählt werden. Im Falle der Abschätzung nach unten wird der größere, im Falle der Abschätzung nach oben der kleinere der beiden berechneten Werte verwendet.

Durch die Wahl von $Z = \frac{\bar{X}}{E(\bar{X})}$ lassen sich damit die gesuchten Abschätzungen der Verteilungsfunktionen von Z und $\bar{X}$ angeben.

5.2. Die Abschätzung der Verteilungsfunktion von Q^a

Im folgenden soll gezeigt werden, daß die Verteilungsfunktion von Q^a , als Funktion von a aufgefaßt, antiton ist, daß also für alle $z \in (0, 1)$

$$\forall a_1, a_2 > 0 \text{ mit } a_2 > a_1 : P_{a_1} \{Q^{a_1} \leq z\} \geq P_{a_2} \{Q^{a_2} \leq z\} \quad (5.13)$$

erfüllt ist.

Aus dem in Kapitel 2.3 hergeleiteten Zusammenhang (2.20) ergibt sich:

Sind $Y_1, \dots, Y_{n-1}$ unabhängige Zufallsvariable mit $Y_i \sim \text{Beta}(a, \frac{1}{n})$, $i = 1, \dots, n-1$, so gilt:

$$\forall a > 0 : P_a \{Q^a \leq z\} = P_a \{Q^{an} \leq z^{an}\} = P_a \left\{ \prod_{i=1}^{n-1} Y_i^a \leq z^n \right\} \quad (5.14)$$

In einem ersten Schritt soll eine zu (5.13) analoge Aussage für die Verteilungsfunktion von Y_i^a hergeleitet werden, womit im Sinne eines Induktionsanfangs die Behauptung (5.13) für den Stichprobenumfang $n = 2$ gezeigt wäre.

Induktionsanfang

Sei $Y \sim \text{Beta}(a, c)$ mit $a > 0$ und $c \in (0, 1)$. Die Dichtefunktion f_a der Transformation $Z \stackrel{\text{def}}{=} Y^a$ lautet für $z \in (0, 1)$:

$$f_a(z) = \frac{1}{aB(a, c)} \left(1 - z^{\frac{1}{a}}\right)^{c-1}$$

Für $\epsilon > 1$ wird die Ableitung des Dichtequotienten

$$q_\epsilon(z) \stackrel{\text{def}}{=} \frac{f_{a\epsilon}(z)}{f_a(z)}$$

nach z betrachtet:

$$\begin{aligned} \frac{\partial}{\partial z} q_\epsilon(z) &= \frac{\partial}{\partial z} z^{\frac{1}{a}} \frac{\partial}{\partial z^{\frac{1}{a}}} q_\epsilon(z) \\ &= \frac{1}{a} z^{\frac{1}{a}-1} \frac{B(a, c)}{\epsilon B(a\epsilon, c)} \left(\frac{1 - z^{\frac{1}{a\epsilon}}}{1 - z^{\frac{1}{a}}} \right)^{c-2} \frac{1 - c}{a z^{\frac{1}{a}} \left(1 - z^{\frac{1}{a}}\right)^2} \\ &\quad \cdot \left\{ \left(1 - z^{\frac{1}{a}}\right) \frac{1}{\epsilon} \left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}-1} - \left(1 - \left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}}\right) \right\} \end{aligned}$$

Die Taylorentwicklung der Funktion $\left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}}$ an der Stelle 1

$$\left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}} = 1 + \left(1 - z^{\frac{1}{a}}\right) \frac{1}{\epsilon} \left(z^{\frac{1}{a}} + \vartheta\right)^{\frac{1}{\epsilon}-1} \quad \text{mit } \vartheta \in \left[0, \left(1 - z^{\frac{1}{a}}\right)\right]$$

führt zu

$$\begin{aligned} \left\{ \left(1 - z^{\frac{1}{a}}\right) \frac{1}{\epsilon} \left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}-1} - \left(1 - \left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}}\right) \right\} = \\ = \left(1 - z^{\frac{1}{a}}\right) \frac{1}{\epsilon} \left\{ \left(z^{\frac{1}{a}}\right)^{\frac{1}{\epsilon}-1} - \left(z^{\frac{1}{a}} + \vartheta\right)^{\frac{1}{\epsilon}-1} \right\} \end{aligned}$$

und damit zur für alle $z \in (0, 1)$ gültigen Ungleichung

$$\forall a > 0 \quad \forall \epsilon \in (0, 1) \quad \forall \epsilon > 1: \quad \frac{\partial}{\partial z} q_{\epsilon}(z) > 0. \quad (5.15)$$

Aufgrund der Normierungsbedingung

$$\int_0^1 f_{a\epsilon}(y) dy = 1 = \int_0^1 f_a(y) dy,$$

existiert somit ein $z_0 \in (0, 1)$, so daß

$$q_{\epsilon}(z) > 1 \quad \Longleftrightarrow \quad z > z_0$$

gilt. z_0 ist damit Nullstelle der für $z \in (0, 1)$ isotonen Funktion $f_{a\epsilon} - f_a$. Für $z \leq z_0$ folgt

$$\int_0^z f_{a\epsilon}(y) dy < \int_0^z f_a(y) dy$$

und für $z > z_0$

$$\int_0^z f_{a\epsilon}(y) dy = 1 - \int_z^1 f_{a\epsilon}(y) dy < 1 - \int_z^1 f_a(y) dy = \int_0^z f_a(y) dy,$$

womit für alle $z \in (0, 1)$ die Behauptung

$$\forall a_1, a_2 > 0 \quad \text{mit } a_2 > a_1: \quad P_{a_1} \{Y_i^{a_1} \leq z\} \geq P_{a_2} \{Y_i^{a_2} \leq z\} \quad (5.16)$$

gezeigt ist, da für $i = 1, \dots, n-1$ der Parameter $c = \frac{1}{n}$ stets kleiner 1 ist.

Da für sämtliche Zufallsvariablen $Y_1, \dots, Y_{n-1}$ die gesuchte Monotonieeigenschaft erfüllt ist, bleibt nur noch die naheliegende Behauptung zu überprüfen, daß sie auch für das Produkt der Y_i gilt. Dies kann per Induktion geschehen. Die Formel (5.16) ist für $n = 2$ aufgrund von (5.14) äquivalent zu (5.13), (5.13) ist damit für $n = 2$ bewiesen.

Induktionsschritt

Seien für $n \geq 3$ $Y_1, \dots, Y_{n-1}$ unabhängige Zufallsvariable, und es gelte $Y_i \sim \text{Beta}(a, c_i)$ mit $a > 0$ und $c_i \in (0, 1)$ für $i = 1, \dots, n-1$. Sei für die Verteilungsfunktion $F_a^{Z_1}$ der Zufallsvariablen

$$Z_1 \stackrel{\text{def}}{=} \prod_{i=1}^{n-2} Y_i^a$$

für alle $z \in (0, 1)$ die Monotonieeigenschaft

$$\forall a_1, a_2 > 0 \text{ mit } a_2 > a_1 : F_{a_1}^{Z_1}(z) \geq F_{a_2}^{Z_1}(z) \quad (5.17)$$

bewiesen. Die Verteilungsfunktion $F_a^{Z_2}$ der Zufallsvariablen

$$Z_2 \stackrel{\text{def}}{=} Y_{n-1}^a$$

erfüllt — wie oben gezeigt — die Monotonieeigenschaft (5.16), ihre Dichtefunktion soll mit $f_a^{Z_2}$ bezeichnet werden. Die Verteilungsfunktion des Produktes $Z_1 Z_2$ lautet für $z \in (0, 1)$

$$P_a \{Z_1 Z_2 \leq z\} = P_a \left\{ Z_1 \leq \frac{z}{Z_2} \right\} = \int_0^1 F_a^{Z_1} \left(\frac{z}{y} \right) f_a^{Z_2}(y) dy.$$

Als Funktion von y aufgefaßt, besitzt $F_a^{Z_1} \left(\frac{z}{y} \right)$ im Bereich $y \in (0, z)$ den Wert 1 und fällt monoton im Bereich $y \in [z, 1)$ auf den Wert $F_a^{Z_1}(z)$.

Für $\epsilon > 1$ ergibt sich aus (5.17) die für $z \in (0, 1)$ gültige Ungleichung

$\forall a > 0 \forall \epsilon > 1 :$

$$P_{a\epsilon} \{Z_1 Z_2 \leq z\} - P_a \{Z_1 Z_2 \leq z\} \leq \int_0^1 F_a^{Z_1} \left(\frac{z}{y} \right) f_{a\epsilon}^{Z_2}(y) dy - \int_0^1 F_a^{Z_1} \left(\frac{z}{y} \right) f_a^{Z_2}(y) dy,$$

was unter Einführung von

$$\begin{aligned} r(y) &\stackrel{\text{def}}{=} F_a^{Z_1} \left(\frac{z}{y} \right) \\ s(y) &\stackrel{\text{def}}{=} f_{a\epsilon}^{Z_2}(y) - f_a^{Z_2}(y) \end{aligned}$$

kurz auch mit

$$\forall a > 0 \forall \epsilon > 1 : P_{a\epsilon} \{Z_1 Z_2 \leq z\} - P_a \{Z_1 Z_2 \leq z\} \leq \int_0^1 r(y) s(y) dy$$

bezeichnet wird. Wie bereits gezeigt, ist die Funktion s isoton mit einer Nullstelle $z_0 \in (0, 1)$, womit sich aufgrund der Antitonie der nichtnegativen Funktion r die Ungleichungen

$$\int_0^{z_0} r(y) s(y) dy \leq r(z_0) \int_0^{z_0} s(y) dy$$

und

$$\int_{z_0}^1 r(y) s(y) dy \leq r(z_0) \int_{z_0}^1 s(y) dy$$

ergeben und wegen $\int_0^1 s(y) dy = 0$ gilt schließlich für alle $z \in (0, 1)$ die Ungleichung

$$\forall a > 0 \quad \forall \epsilon > 1 : \quad P_a \{Z_1 Z_2 \leq z\} \geq P_{a\epsilon} \{Z_1 Z_2 \leq z\}.$$

Damit ist die Behauptung (5.17) auch für die Verteilungsfunktion des Produkts $Z_1 Z_2$ gezeigt und damit auch die Behauptung (5.13).

5.3. Die Abschätzung der Überdeckungswahrscheinlichkeiten

Die Abschätzung der Überdeckungswahrscheinlichkeiten kann mit Hilfe der in den Abschnitten 5.1 und 5.2 gewonnenen Abschätzungen der Verteilungsfunktionen von $\frac{\bar{X}}{E(\bar{X})}$ und Q^a vorgenommen werden. Wie zu Beginn dieses Kapitels dargestellt, werden, für ein durch eine Funktion c gegebenes Konfidenzintervall $C(\bar{X})$, je eine obere und untere Schranke der Funktion $\tilde{\alpha}$ gesucht. Diese Abschätzungen sollen für jedes beliebige Intervall (a_1, a_2) , mit $0 < a_1 < a_2$ berechenbar sein. Der Funktionswert von $\tilde{\alpha}$ an der Stelle a läßt sich in der folgenden Weise darstellen:

$$\begin{aligned} \tilde{\alpha}(a) &= P_{a, \sigma^2} \left\{ \sigma^2 > c(\hat{a}_{ML}) \hat{\sigma}_{ML}^2 \right\} \\ &= P_{a, \sigma^2} \left\{ \bar{X} < \sqrt{\sigma^2 \frac{\hat{a}_{ML}}{c(\hat{a}_{ML})}} \right\} \\ &= P_{a, \sigma^2} \left\{ \frac{\bar{X}}{E(\bar{X})} < \sqrt{\frac{\hat{a}_{ML}}{ac(\hat{a}_{ML})}} \right\}. \end{aligned} \quad (5.18)$$

Mit Hilfe der in Kapitel 2.4 eingeführten Funktionen ρ^{-1} und τ kann wegen

$$\hat{a}_{ML} = \rho^{-1}(\ln Q) = \frac{\tau(\ln Q)}{\ln Q}$$

(5.18) weiter umgeformt werden:

$$\begin{aligned}\tilde{\alpha}(a) &= P_{a, \sigma^2} \left\{ \frac{\bar{X}}{E(\bar{X})} < \sqrt{\frac{1}{-a \ln Q}} \sqrt{\frac{-\tau(\ln Q)}{c \circ \rho^{-1}(\ln Q)}} \right\} \\ &= \int_0^1 \mathcal{G}_{a n, \frac{1}{a n}} \left(\sqrt{\frac{1}{-a \ln q}} \sqrt{\frac{-\tau(\ln q)}{c \circ \rho^{-1}(\ln q)}} \right) dF_{a, n}^Q(q),\end{aligned}\quad (5.19)$$

wobei mit $F_{a, n}^Q$ die Verteilungsfunktion von Q bezeichnet werden soll.

Im folgenden werden einfache Abschätzungen der einzelnen „Komponenten“ der Formel (5.19) gegeben, die es ermöglichen, das gesamte Integral nach unten und oben abzuschätzen. Dabei werden die Abschätzungen in zweierlei Hinsicht vorgenommen werden müssen. Einerseits werden Abschätzungen für die Variation des Parameters a im Intervall (a_1, a_2) benötigt, andererseits können die in (5.19) angegebenen Größen auch nur für endlich viele Werte von q bestimmt werden, so daß auch der sich daraus ergebende Effekt berücksichtigt werden muß.

Abschätzung des Arguments von $\mathcal{G}_{a n, \frac{1}{a n}}$

In einem ersten Schritt soll die im Argument von $\mathcal{G}_{a n, \frac{1}{a n}}$ vorkommende Größe

$\sqrt{\frac{-\tau(\ln q)}{c \circ \rho^{-1}(\ln q)}}$ für $q \in (q_1, q_2)$, $0 < q_1 < q_2 < 1$ betrachtet werden. Wie in Kapitel 4 erläutert, wird die Funktion c im Stützstellenbereich $\hat{A}$ numerisch bestimmt und mit Hilfe der in Kapitel 4.9 vorgestellten Interpolationsvorschrift auf dem gesamten Bereich der positiven reellen Zahlen definiert. Die Interpolationsvorschrift ist so beschaffen, daß die Funktion c außerhalb des Stützstellenbereiches $\hat{A}$ keine Minima und Maxima besitzen kann (mit Ausnahme eventueller Extremwerte für $\hat{a} \rightarrow 0$ oder $\hat{a} \rightarrow \infty$). Damit ist es auf einfache Weise möglich, für beliebige Intervalle $(\hat{a}_1, \hat{a}_2)$, $0 < \hat{a}_1 < \hat{a}_2 < \infty$

$$\inf_{\hat{a}_1 < \hat{a} < \hat{a}_2} c(\hat{a}) \quad \text{und} \quad \sup_{\hat{a}_1 < \hat{a} < \hat{a}_2} c(\hat{a})$$

anzugeben. In allen Beispielen, die berechnet wurden, war c zudem im gesamten Stützstellenbereich und damit auch im ganzen Raum $\mathbb{R}^+$ streng antiton, so daß im allgemeinen $\inf_{\hat{a}_1 < \hat{a} < \hat{a}_2} c(\hat{a}) = c(\hat{a}_2)$ und $\sup_{\hat{a}_1 < \hat{a} < \hat{a}_2} c(\hat{a}) = c(\hat{a}_1)$ gilt. Da die in Kapitel 2.4 betrachtete Funktion ρ^{-1} streng isoton ist, können somit auch die beiden Größen $\inf_{q_1 < q < q_2} c \circ \rho^{-1}(\ln q)$ und $\sup_{q_1 < q < q_2} c \circ \rho^{-1}(\ln q)$ bestimmt werden, und es wird im allgemeinen $\inf_{q_1 < q < q_2} c \circ \rho^{-1}(\ln q) = c \circ \rho^{-1}(\ln q_2)$ sowie $\sup_{q_1 < q < q_2} c \circ \rho^{-1}(\ln q) = c \circ \rho^{-1}(\ln q_1)$ gelten. Die ebenfalls in Kapitel 2.4 eingeführte Funktion τ ist streng isoton mit den Grenzwerten

$$\lim_{q \rightarrow 0} \tau(\ln q) = -1 \quad \text{und} \quad \lim_{q \rightarrow 1} \tau(\ln q) = -\frac{1}{2}.$$

Im Intervall (q_1, q_2) ergibt sich damit:

$$\inf_{q_1 < q < q_2} -\tau(\ln q) = -\tau(\ln q_2) \quad \text{und} \quad \sup_{q_1 < q < q_2} -\tau(\ln q) = -\tau(\ln q_1).$$

Dieses ermöglicht bereits die Abschätzung

$$\sqrt{\frac{-\tau(\ln q_2)}{\sup_{q_1 < q < q_2} c \circ \rho^{-1}(\ln q)}} < \sqrt{\frac{-\tau(\ln q)}{c \circ \rho^{-1}(\ln q)}} < \sqrt{\frac{-\tau(\ln q_1)}{\inf_{q_1 < q < q_2} c \circ \rho^{-1}(\ln q)}}. \quad (5.20)$$

Im allgemeinen wird aber nicht nur die Antitonie von c gelten, selbst der Quotient $\frac{c \circ \rho^{-1}(\ln q)}{\tau(\ln q)}$ wird aufgrund der geringen Steigung der beschränkten Funktion $-\tau \circ \ln$ noch monoton fallend sein, so daß sich für alle $q \in (q_1, q_2)$ die schärfere Abschätzung

$$\sqrt{\frac{-\tau(\ln q_1)}{c \circ \rho^{-1}(\ln q_1)}} < \sqrt{\frac{-\tau(\ln q)}{c \circ \rho^{-1}(\ln q)}} < \sqrt{\frac{-\tau(\ln q_2)}{c \circ \rho^{-1}(\ln q_2)}} \quad (5.21)$$

verifizieren läßt (dieses muß natürlich im Einzelfall nachgeprüft werden).

Die Einführung der positiven Zufallsvariablen

$$Z \stackrel{\text{def}}{=} -a \ln Q,$$

deren Verteilungsfunktion mit $F_{a,n}^Z$ bezeichnet werden soll, erlaubt eine weitere Umformung von (5.19):

$$\tilde{\alpha}(a) = \int_0^1 \mathcal{G}_{an, \frac{1}{an}} \left(\frac{1}{\sqrt{z}} \sqrt{\frac{-\tau\left(-\frac{z}{a}\right)}{c \circ \rho^{-1}\left(-\frac{z}{a}\right)}} \right) dF_{a,n}^Z(z). \quad (5.22)$$

In dieser Form soll nun eine für alle $a \in (a_1, a_2)$ gültige Abschätzung des Integrals

$$\int_{z_1}^{z_2} \mathcal{G}_{an, \frac{1}{an}} \left(\frac{1}{\sqrt{z}} \sqrt{\frac{-\tau\left(-\frac{z}{a}\right)}{c \circ \rho^{-1}\left(-\frac{z}{a}\right)}} \right) dF_{a,n}^Z(z) \quad (5.23)$$

gegeben werden. Für $z \in (z_1, z_2)$ und $a \in (a_1, a_2)$ gilt

$$-\frac{z_2}{a_1} < -\frac{z}{a} < -\frac{z_1}{a_2},$$

woraus sich mit $q_1 \stackrel{\text{def}}{=} e^{-z_2/a_1}$ und $q_2 \stackrel{\text{def}}{=} e^{-z_1/a_2}$ Abschätzungen von $\sqrt{\frac{-\tau\left(-\frac{z}{a}\right)}{c \circ \rho^{-1}\left(-\frac{z}{a}\right)}}$ durch Einsetzen in (5.20) bzw. (5.21) ergeben. Die untere Schranke soll mit η , die obere Schranke mit $\bar{\eta}$ bezeichnet werden. Für das Argument von $\mathcal{G}_{an, \frac{1}{an}}$ gilt damit die Abschätzung

$$\underline{\eta} \sqrt{\frac{1}{z_2}} < \sqrt{\frac{1}{z}} \sqrt{\frac{-\tau\left(-\frac{z}{a}\right)}{c \circ \rho^{-1}\left(-\frac{z}{a}\right)}} < \bar{\eta} \sqrt{\frac{1}{z_1}}.$$

Aus den in Abschnitt 5.1 hergeleiteten unteren und oberen Schranken $\underline{\mathcal{G}}_{an, \frac{1}{an}}$ und $\overline{\mathcal{G}}_{an, \frac{1}{an}}$ der Verteilungsfunktion $\mathcal{G}_{an, \frac{1}{an}}$ ergibt sich somit

$$\begin{aligned} \underline{\mathcal{G}}_{an, \frac{1}{an}} \left(\eta \sqrt{\frac{1}{z_2}} \right) & \inf_{a_1 < a < a_2} (F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)) \\ & < \int_{z_1}^{z_2} \underline{\mathcal{G}}_{an, \frac{1}{an}} \left(\frac{1}{\sqrt{z}} \sqrt{\frac{-\tau(-\frac{z}{a})}{c \circ \rho^{-1}(-\frac{z}{a})}} \right) dF_{a,n}^Z(z) \quad (5.24) \\ & < \overline{\mathcal{G}}_{an, \frac{1}{an}} \left(\bar{\eta} \sqrt{\frac{1}{z_1}} \right) \sup_{a_1 < a < a_2} (F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)). \end{aligned}$$

Die Abschätzung der Differenz $F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)$

Die Schranken des Integrals (5.23) sind gefunden, wenn es gelingt, für $a \in (a_1, a_2)$ gültige Abschätzungen der Differenz $F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)$ anzugeben. Dies ist aber mit Hilfe der in Abschnitt 5.2 gezeigten Monotonieeigenschaft (5.13) der Verteilungsfunktion von Q^a leicht möglich, wie aus der Umformung

$$F_{a,n}^Z(z) = P_a \{-a \ln Q \leq z\} = P_a \{Q^a \geq e^{-z}\}$$

ersichtlich ist. Aus (5.13) folgt nun unmittelbar die Gültigkeit von

$$\inf_{a_1 < a < a_2} F_{a,n}^Z(z) = F_{a_1,n}^Z(z) \quad \text{und} \quad \sup_{a_1 < a < a_2} F_{a,n}^Z(z) = F_{a_2,n}^Z(z).$$

Für die Differenz $F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)$ kann nun die einfache Abschätzung

$$F_{a_1,n}^Z(z_2) - F_{a_2,n}^Z(z_1) < F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1) < F_{a_2,n}^Z(z_2) - F_{a_1,n}^Z(z_1)$$

angegeben werden, die sich aufgrund des Zusammenhangs $F_{a,n}^Z(z) = 1 - F_{a,n}^Q(e^{-\frac{z}{a}})$ auch in der Form

$$F_{a_2,n}^Q(e^{-\frac{z_2}{a_2}}) - F_{a_1,n}^Q(e^{-\frac{z_1}{a_1}}) < F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1) < F_{a_1,n}^Q(e^{-\frac{z_1}{a_1}}) - F_{a_2,n}^Q(e^{-\frac{z_2}{a_2}}) \quad (5.25)$$

darstellen läßt.

Die Abschätzung von $\tilde{\alpha}$

Für jede beliebige Zerlegung $0 < z_1 < z_2 < \dots < z_k < \infty$ lassen sich für $\tilde{\alpha}$ damit für alle $a \in (a_1, a_2)$ gültige Schranken angeben, indem die sich aus (5.24) und (5.25) ergebenden Abschätzungen für jedes Teilintervall (z_j, z_{j+1}) , $j = 1, \dots, k-1$, sowie die Intervalle $(0, z_1)$ und (z_k, ∞) ermittelt und summiert werden. Für die letzten beiden Intervalle gelten dabei die Schranken

$$\underline{\mathcal{G}}_{an, \frac{1}{an}} \left(\eta \sqrt{\frac{1}{z_1}} \right) F_{a_1,n}^Z(z_1) < \int_0^{z_1} \underline{\mathcal{G}}_{an, \frac{1}{an}} \left(\frac{1}{\sqrt{z}} \sqrt{\frac{-\tau(-\frac{z}{a})}{c \circ \rho^{-1}(-\frac{z}{a})}} \right) dF_{a,n}^Z(z) < F_{a_2,n}^Z(z_1)$$

und

$$0 < \int_{z_k}^{\infty} \underline{\mathcal{G}}_{an, \frac{1}{an}} \left(\frac{1}{\sqrt{z}} \sqrt{\frac{-\tau(-\frac{z}{a})}{c \circ \rho^{-1}(-\frac{z}{a})}} \right) dF_{a,n}^Z(z) < \overline{\mathcal{G}}_{an, \frac{1}{an}} \left(\bar{\eta} \sqrt{\frac{1}{z_k}} \right) (1 - F_{a_1,n}^Z(z_k)).$$

Zur Güte der Abschätzungen

Im nächsten Abschnitt wird an einigen Beispielen demonstriert, daß es mit dem vorliegenden Verfahren möglich ist, durchaus akzeptable Schranken für die Funktion $\tilde{\alpha}$ anzugeben. Die in diesem Abschnitt vorgeschlagenen Abschätzungen, die zum großen Teil auf Monotonieeigenschaften beruhen, erscheinen zwar zum Teil recht grob, insbesondere die Abschätzung der Differenz $F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)$, doch hat der Parameterwert von a , wie an früherer Stelle schon festgestellt wurde, keine allzu starken Auswirkungen auf den Wert der Verteilungsfunktion $F_{a,n}^Z$, so daß die Abschätzungen dennoch recht brauchbar sind. Im nächsten Abschnitt soll auch kurz auf die Wahl der Stützstellen $z_1, \dots, z_k$ eingegangen werden.

5.4. Einige Beispiele

Soll das oben beschriebene Verfahren angewendet werden, stellt sich natürlich die Frage einer geeigneten Wahl der Stützstellen $z_1, \dots, z_k$ und der Zerlegung des Parameterraums in endlich viele Intervalle (a_j, a_{j+1}) , $j = 1, \dots, m$. Die den folgenden Beispielen zugrundeliegenden Aufteilungen sind sicherlich nicht ideal, doch soll auch nur demonstriert werden, daß das Verfahren grundsätzlich zur Abschätzung der Funktion $\tilde{\alpha}$ geeignet ist. Die Berechnungen wurden dabei an einem Personal Computer durchgeführt, die Übertragung auf einen modernen Großrechner kann ohne weiteres einen Geschwindigkeitszuwachs um den Faktor 50 (und mehr) bringen, so daß die Stützstellenzahl an solchen Rechnern bei gleichem Zeitaufwand um ein Vielfaches größer gewählt werden kann als in den vorliegenden Beispielen. Zusammen mit einer durchaus möglichen Optimierung der Stützstellenwahl liegen damit, was die Steigerung der Genauigkeit betrifft, im Vergleich mit den im folgenden vorgestellten Abschätzungen noch beträchtliche „Reserven“ in dem Verfahren.

Zur Wahl der Stützstellen $z_1, \dots, z_k$

Ein erster Anhaltspunkt, wie die Stützstellen $z_1, \dots, z_k$ gewählt werden können, ergibt sich aus den in Kapitel 4.3 vorgestellten Überlegungen zur Wahl der Stützstellen bei der Integration der Funktion h_a . Ein einfacher Ansatz wäre es, die dort diskutierte Stützstellenmenge $u_1, \dots, u_k$ einfach zu übertragen. Aus der in Kapitel 4.2 eingeführten Substitution $u = \mathcal{A}_{a,n}(\hat{a})$ erlaubt der Zusammenhang

$$z = -a \ln q = -a \rho(\mathcal{A}_{a,n}^{-1}(u))$$

damit die direkte Transformation dieser Stützstellenmenge. Werden für die Funktion $\tilde{\alpha}$ Schranken in einem Intervall (a_1, a_2) gesucht und ist dieses Intervall sehr kurz, so gewährleistet diese Wahl, die für ein festes $a \in (a_1, a_2)$ durchgeführt werden muß, also z.B.

$$z_j \stackrel{\text{def}}{=} -a_1 \rho(\mathcal{A}_{a_1,n}^{-1}(u_j)), \quad j = 1, \dots, k, \quad (5.26)$$

die Übertragung der in Kapitel 4.3 vorgestellten Überlegungen zur „Güte“ der Stützstellenwahl von $u_1, \dots, u_k$ auch auf diese allgemeinere Situation. Wird in

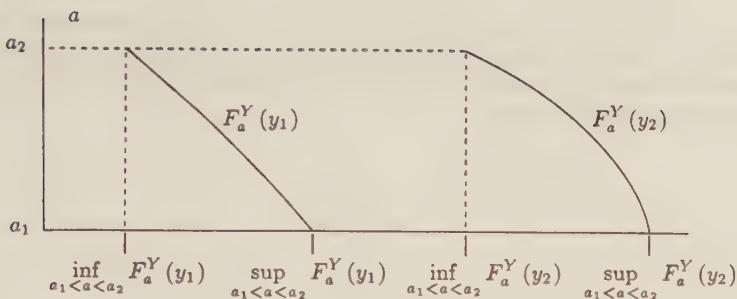


Abb. 5.2 Die Abschätzung von $F_a^Y(y_2) - F_a^Y(y_1)$

dem Verfahren, welches in den ersten drei Abschnitten dieses Kapitels beschrieben wurde, $a_1 = a_2$ gesetzt, so ergeben sich als obere und untere Schranken die im Sinne des Stieltjes-Integralbegriffes definierten Ober- und Untersummen zur gewählten Zerlegung $u_1, \dots, u_k$ bzw. $z_1, \dots, z_k$. In den Abbildungen 5.4–5.6 werden die sich für die in Abbildung 4.5 angegebene Stützstellenmenge $u_1, \dots, u_{238}$ ergebenden Obersummen stets mit $\hat{\alpha}^2$, die Untersummen mit $\hat{\alpha}_2$ bezeichnet. Sie werden im Bereich der Stützstellenmenge A berechnet. Da für eine Stieltjes-integrierbare Funktion — und h_a ist integrierbar — eine Verfeinerung der Zerlegung $u_1, \dots, u_k$ im allgemeinen eine geringere Differenz von Ober- und Untersummen zur Folge hat, wird dieser Effekt anhand der Ober- und Untersummen $\hat{\alpha}^4$ und $\hat{\alpha}_4$, die sich aus einer auf etwa 2000 vergrößerten Stützstellenzahl ergeben, anhand des Beispiels $n = 50$, $1 - \alpha = 0,95$ demonstriert (Abb. 5.6). Diese beiden Funktionen zeigen, daß das Konfidenzintervall so gewählt wurde, daß die Funktion $\bar{\alpha}$ zumindest im Stützstellenbereich A auch unter Berücksichtigung der maximal möglichen Integrationsfehler bereits sehr nahe an der Konstanten α liegt.

Die so gewonnene Zerlegung kann aber für $a_2 > a_1$ — und diese Situation soll ja gerade betrachtet werden — zu fein sein. Die Ursache hierfür ist die in (5.25) angegebene Abschätzung der Differenz $F_{a,n}^Z(z_2) - F_{a,n}^Z(z_1)$. Für die in Abschnitt 5.2 diskutierte Zufallsvariable $Y \stackrel{\text{def}}{=} Q^a$ und $y_1 \stackrel{\text{def}}{=} e^{-z_2}$ sowie $y_2 \stackrel{\text{def}}{=} e^{-z_1}$ gilt es, für $a \in (a_1, a_2)$ gültige Schranken der Differenz $F_a^Y(y_2) - F_a^Y(y_1)$ zu finden. Die obere Schranke wird gemäß (5.25) durch die Differenz einer Abschätzung von $F_a^Y(y_2)$ nach oben und einer Abschätzung von $F_a^Y(y_1)$ nach unten bestimmt, für die untere Schranke werden beide Größen in der umgekehrten Richtung abgeschätzt, wobei die in Abschnitt 5.2 angegebene Monotoniebeziehung (5.13)

$$\begin{aligned} F_{a_2}^Y(y_2) - F_{a_1}^Y(y_1) &= \inf_{a_1 < a < a_2} F_a^Y(y_2) - \sup_{a_1 < a < a_2} F_a^Y(y_1) \\ &< F_a^Y(y_2) - F_a^Y(y_1) \\ &< \sup_{a_1 < a < a_2} F_a^Y(y_2) - \inf_{a_1 < a < a_2} F_a^Y(y_1) = F_{a_1}^Y(y_2) - F_{a_2}^Y(y_1) \end{aligned}$$

in Abb. 5.2 illustriert wird.

In Abb. 5.3 wird verdeutlicht, was passiert, wenn die Differenz zwischen y_1 und y_2 verringert wird. Die Abschätzung nach unten liefert einen trivialen negativen Wert, die Abschätzung nach oben beträgt ein Vielfaches des tatsächlichen Supremums der Differenzen.

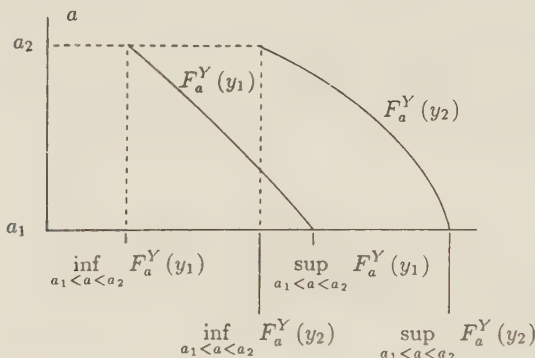


Abb. 5.3 Die Abschätzung von $F_a^Y(y_2) - F_a^Y(y_1)$ bei zu „kleinem“ $y_2 - y_1$

Während die Abschätzung des Arguments der Gammaverteilungsfunktion $\mathcal{G}_{a_n, \frac{1}{\delta_n}}$ mit geringerem Stützstellenabstand also immer schärfer wird, trifft für die Abschätzung der Differenz $F_a^Y(y_2) - F_a^Y(y_1)$ im wesentlichen die umgekehrte Aussage zu. Dieses bedeutet, daß für ein gegebenes Intervall (a_1, a_2) ein Optimierungsproblem entsteht.

Diese Fragestellung soll hier nicht vertieft behandelt werden. Es soll lediglich ein kurzer Einblick in eine prinzipielle Strategie gegeben werden. Für ein gegebenes Intervall (z_1, z_2) seien die Schranken (5.24) gegeben. Das Intervall wird in zwei Teilintervalle (z_1, z_3) und (z_3, z_2) zerlegt. Für jedes der beiden Teilintervalle werden wiederum Abschätzungen gemäß (5.24) gebildet. Ist die Summe der unteren Schranken der beiden Teilintervalle größer als die untere Schranke, die unter Zugrundelegung des gesamten Intervalls berechnet wurde, so war die Zerlegung „erfolgreich“ und die gleiche Strategie wird rekursiv auf die beiden Teilintervalle angewendet, ansonsten war die Zerlegung nicht erfolgreich, und es wird je nach Strategie entweder eine andere Zerlegung des Intervalls (z_1, z_2) im obigen Sinne „getestet“, oder es wird nicht weiter versucht, die Abschätzung zu verbessern. Analog wird die Abschätzung nach oben behandelt, wobei eine Verbesserung sich hier selbstverständlich in einem Absinken der Schranke ausdrückt. Ein Kriterium dafür, ob eine weitere Zerlegung getestet wird, kann zum Beispiel die für ein gegebenes Intervall bereits erreichte Genauigkeit sein, die sich durch die Differenz der oberen und unteren Schranke ausdrücken läßt. Das Verfahren beginnt mit einer Zerlegung des gesamten Intervalls $(0, 1)$, für welches trivialerweise die Schranken 0 und 1 existieren.

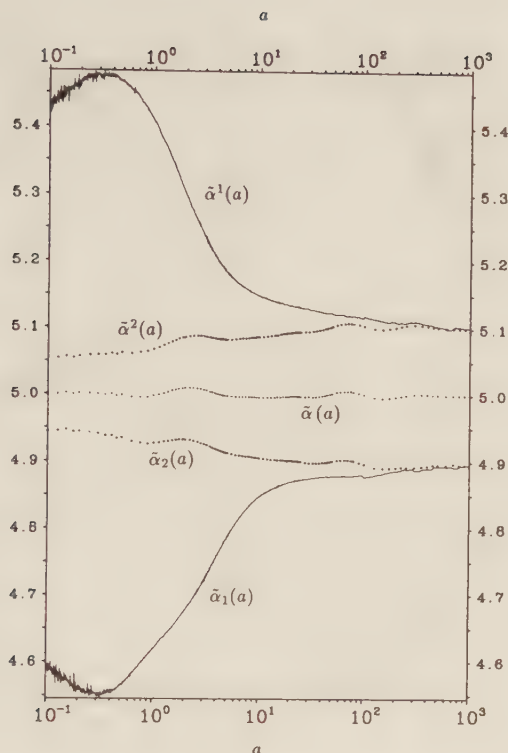


Abb. 5.4 Die Abschätzungen für das Beispiel $n = 10$

Beispiele

In den bereits angekündigten Abbildungen 5.4–5.6 sind für $1 - \alpha = 0,95$ und die Stichprobenumfänge $n = 10$ (Abb. 5.4), $n = 250$ (Abb. 5.5) und $n = 50$ (Abb. 5.6) untere und obere Schranken angegeben. Das Verfahren zur Bestimmung der Zerlegung wurde so gewählt, daß in jedem Falle nur die Stützstellen $z_1, \dots, z_{238}$, die über (5.26) aus der in Abb. 4.5 angegebenen Stützstellenmenge $u_1, \dots, u_{238}$ definiert sind, zur Berechnung verwendet wurden. Gemäß dem angewendeten Zerlegungsalgorithmus wurde im allgemeinen eine weniger feine Zerlegung errechnet, auf keinen Fall aber eine feinere. Die oben eingeführten Funktionen $\tilde{\alpha}_2$ und $\tilde{\alpha}^2$ stellen damit Unter- bzw. Obergrenzen für die erreichbare Abschätzungsgenauigkeit dar. Für die gewählte Zerlegung des Parameterraumes von etwa 800 Stützstellen — wobei im Bereich kleiner Parameterwerte eine erheblich feinere Zerlegung gewählt wurde als für große (in dem Sinne, daß der

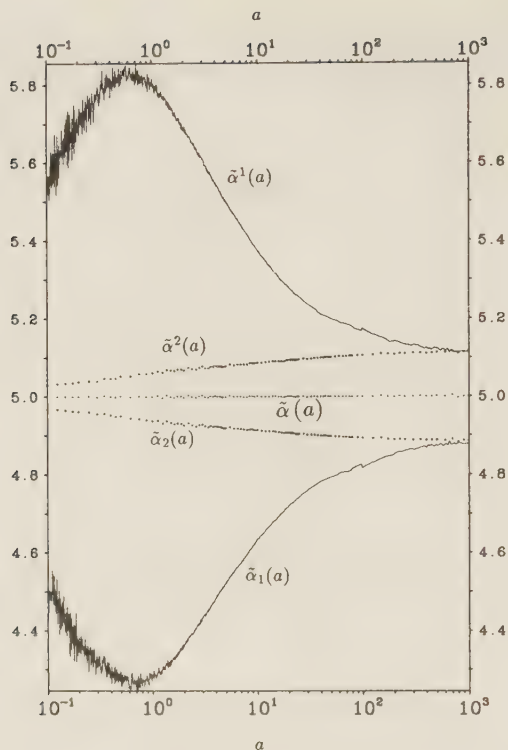


Abb. 5.5 Die Abschätzungen für das Beispiel $n = 250$

Quotient aus oberer und unterer Grenze a_2 und a_1 wesentlich kleiner war) — ergaben sich als untere und obere Schranken die mit $\tilde{\alpha}_1$ und $\tilde{\alpha}^1$ bezeichneten Kurven. Die Schwankungen der Funktionen lassen sich dadurch erklären, daß die gewählte Zerlegung für jedes Intervall neu berechnet wurde. Insbesondere im Bereich kleiner Werte von a , bei denen zum Teil nur sehr wenige Intervalle zur Berechnung herangezogen wurden, machen sich Änderungen der Zerlegung deutlich bemerkbar. Die Vorgabe der festen Stützstellenmenge $u_1, \dots, u_{238}$, die bei gleichem Rechenzeitaufwand eine feinere Zerlegung des Parameterraums erlaubt, wirkt sich für kleine und mittlere Werte des Parameters nicht nachteilig auf die Genauigkeit der Schranken $\tilde{\alpha}_1$ und $\tilde{\alpha}^1$ aus, im Gegenteil, in diesem Bereich kommt es im wesentlichen auf eine möglichst feine Zerlegung des Parameterraums an. Im Bereich großer Werte von a „verhinderte“ — wie durch den Vergleich der Funktionen $\tilde{\alpha}_1, \tilde{\alpha}^1$ und $\tilde{\alpha}_2, \tilde{\alpha}^2$ sowie $\tilde{\alpha}_4, \tilde{\alpha}^4$ offenkundig wird — diese Vorgabe

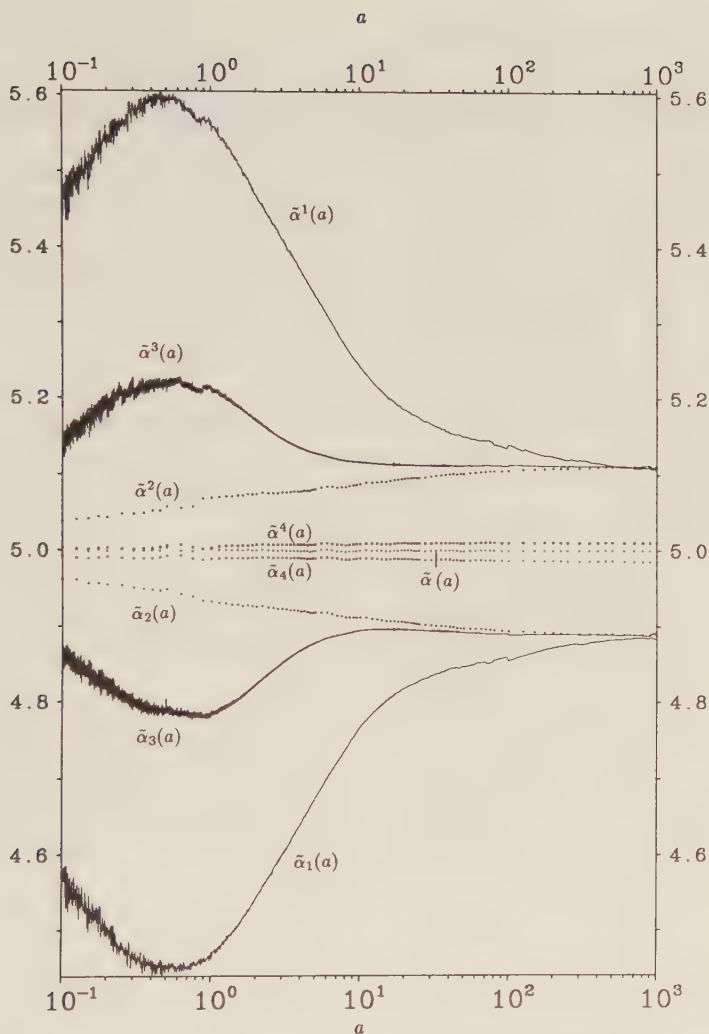


Abb. 5.6 Die Abschätzungen für das Beispiel $n = 50$

jedoch bessere Abschätzungen. Für das Beispiel $n = 50$ sind mit $\tilde{\alpha}_3$ und $\tilde{\alpha}^3$ noch diejenigen Schranken angegeben, die sich durch eine um den Faktor 7 verfeinerte Zerlegung des Parameterraumes ergeben. Deutlich zu erkennen ist die verbesserte Annäherung an die Funktionen $\tilde{\alpha}_2$ und $\tilde{\alpha}^2$.

6. Zusammenfassung

Die vorliegende Arbeit hatte zum Ziel, ein einseitiges Konfidenzintervall für die Varianz der Gammaverteilung zu konstruieren. Die Familie der Gammaverteilungen stellt dabei eine sehr große Klasse von rechtsschiefen Verteilungen dar. Der zweite Parameter der Gammaverteilung, der Gestaltparameter, wurde als unbekannt angenommen. Optimal in dieser Situation wäre es, ein ähnliches Konfidenzintervall in bezug auf diesen unbekannten Parameter angeben zu können. Da die Frage der Existenz eines solchen Konfidenzintervalls nicht geklärt ist, wird in Kapitel 4 ein konstruktives Verfahren entwickelt, welches dieser Idealvorstellung möglichst gut entsprechen soll. Die numerischen Untersuchungen zeigen, daß dieser Anspruch für eine große Anzahl von Werten des Parameters α — wobei die Zahl im wesentlichen durch die Kapazitätsgrenzen des verwendeten Rechners beschränkt ist — bei vertretbarem Rechenaufwand durch das Verfahren eingelöst werden kann.

Die in Kapitel 4.10 vorgestellten Beispiele zur Anwendung des Verfahrens ergeben, daß sich bei der dort vorgestellten Darstellung der Funktion c in der Gestalt $c(\hat{a}) = (c_{n,\alpha})^{1 + \frac{\delta^*(\hat{a})}{4}}$ das Konfidenzniveau und der Stichprobenumfang maßgeblich im Wert der Konstanten $c_{n,\alpha}$ widerspiegeln, der Funktionswert der Funktion δ^* wird von diesen beiden Größen nur relativ geringfügig beeinflusst. Da die Funktion δ^* zudem auch in ihrem Argument $\hat{a}$ nur geringfügig variiert und die Konstante $c_{n,\alpha}$ sich aus Tabellen der χ^2 -Verteilung in einfacher Weise bestimmen läßt, ist eine Darstellung des Konfidenzintervalls gefunden, die eine Tabellierung mit sehr geringem Aufwand erlaubt. Theoretische und weitere praktische Untersuchungen dieser „Auffälligkeiten“ wären wünschenswert, hätten den Rahmen dieser Arbeit jedoch überschritten.

In Kapitel 5 schließlich wird durch Abschätzungen die Einhaltung des Konfidenzniveaus zwischen den Stützstellen kontrolliert. Die befriedigenden Ergebnisse dieser Abschätzungen werden anhand einiger Beispiele illustriert. Zusammenfassend kann festgestellt werden, daß es zumindest im Falle der berechneten Beispiele, die schon einen relativ großen Bereich von Stichprobenumfängen und Konfidenzniveaus abdecken, mit dem Verfahren gelingt, das Konfidenzniveau für praktische Zwecke hinreichend genau einzuhalten.

Symbolverzeichnis

Die angegebene Seitenzahl bezieht sich auf die erstmalige Verwendung des Symbols.

Kapitel 2:

a	4
b	4
$f_{a,b}(\cdot)$	4
$G(a, b)$	4
σ^2	7
$G(a, \sigma^2)$	8
$\mathcal{G}_{a, \sigma^2}(\cdot)$	8
$f_{a, \sigma^2}(\cdot)$	9
$\psi(\cdot)$	11
$\hat{a}_{ML}$	11
$\hat{\sigma}_{ML}^2$	11
Q	11
$f_a^Q(q)$	13
$\rho(\cdot)$	18
$\tau(\cdot)$	21
$f_{a, ML}^{\hat{a}}(\hat{a})$	21
$\psi^{(k)}(\cdot)$	23
$\mathcal{A}_{a,n}(\hat{a})$	29
$\tilde{\mathcal{A}}_{\hat{a},n}(a)$	29
$\mathcal{A}_{a,n}^{-1}(u)$	29
$\tilde{\mathcal{A}}_{\hat{a},n}^{-1}(u)$	29

Kapitel 3:

$w(\cdot)$	33
$p_m(\cdot)$	33
$I_J(\cdot)$	35

Kapitel 4:

$\underline{X}, \underline{x}$	50
$C(\underline{X}), C(\underline{x})$	50

Kapitel 4:

$c(\underline{X}), c(\underline{x})$	50
$1 - \alpha$	50
A	51
$H_{\sigma_0^2}^0, H_{\sigma_0^2}^1$	53
$\Phi_{\sigma_0^2}(\underline{X}), \Phi_{\sigma_0^2}(\underline{x})$	53
$c(\hat{a}_{ML}), c(\hat{a})$	55
$\tilde{\alpha}(a)$	57
$h_a(u)$	57
$c^{(i)}(\hat{a}_{ML}), c^{(i)}(\hat{a})$	58
$\tilde{\alpha}^{(i)}(a)$	58
$\hat{A}$	60
$\{u_1, \dots, u_{238}\}$	61
$C^*(\underline{X}), C^*(\underline{x})$	62
$H_{a_0, \sigma_0^2}^0, H_{a_0, \sigma_0^2}^1$	62
$c^*(a, \hat{a}_{ML}), c^*(a, \hat{a})$	62
$\gamma_{an}(\cdot)$	65
$H_{a_0, \sigma_0^2}^{0*}$	68
U, V	68
$h_a^*(u)$	69
$h^*(u)$	71
$\xi_{\hat{a}, \frac{\sigma^2}{\sigma^2}}(u)$	74
$\xi_{\hat{a}, c(\hat{a})}(u)$	76
$\beta^{(i)}(\hat{a})$	88
c_∞	92
$c_{n, \alpha}$	93
$\delta^*(\hat{a})$	93

Kapitel 5:

$\underline{\eta}, \underline{\bar{\eta}}$	112
$\underline{\mathcal{G}}_{an, \frac{1}{an}}(\cdot), \overline{\mathcal{G}}_{an, \frac{1}{an}}(\cdot)$	113

Literaturverzeichnis

- Abramowitz, M. and I. Stegun (1984): Pocketbook of mathematical functions; abridged edition of: Handbook of mathematical functions
Harri Deutsch, Thun, Frankfurt/Main
- Bain, L. J. and M. Engelhardt (1975): A two-moment chi-square approximation for the statistic $\ln \frac{\bar{x}}{\bar{x}}$
Journal of the American Statistical Society, **70**, 948–950
- Bar-Lev, S. K. and B. Reiser (1982): An exponential subfamily which admits UMPU tests based on a single test statistic
The Annals of Statistics, **10**, 979–989
- Barndorff-Nielsen, O. (1978): Information and exponential families in statistical theory
John Wiley & Sons, Chichester, New York, Brisbane, Toronto
- Basu, D. (1955): On statistics independent of a complete sufficient statistic
Sankhyā, **15**, 377–380
- Berman, M. (1981): The maximum likelihood estimators of the parameters of the gamma distribution are always positively biased
Communications in Statistics, Theory and Methods, **A 10(7)**, 693–697
- Bishop, D. J. and U. S. Nair (1939): A note on certain methods of testing for the homogeneity of a set of estimated variances
Journal of the Royal Statistical Society, Supplement, **6**, No. 1, 89–99
- Bowman, K. O. and L. R. Shenton (1988): Properties of estimators for the gamma distribution
Marcel Dekker, Inc., New York, Basel
- Box, G. E. P. (1949): A general distribution theory of a class of likelihood criteria
Biometrika, **36**, 317–346
- Bronstein, I. N. und K. A. Semendjajew (1987): Taschenbuch der Mathematik
Verlag Nauka, Moskau; BSB B. G. Teubner, Leipzig
- Choi, S. C. and R. Wette (1969): Estimation of the parameters of the gamma distribution
Technometrics, **11**, 683–690
- Cox, D. R. and P. A. W. Lewis (1966): The statistical analysis of series of events
Methuen & Co., London

- Engelhardt, M. and L. J. Bain (1977): Uniformly most powerful unbiased tests on the scale parameter of a gamma distribution with a nuisance shape parameter
Technometrics, **19**, 77–80
- Engelhardt, M. and L. J. Bain (1978): Uniformly most powerful unbiased tests on the scale parameter of the gamma distribution
Technometrics, **20**, 485–489
- Feller, W. (1971): An introduction to probability theory and its application, Vol. II
John Wiley & Sons, New York, Chichester, Brisbane, Toronto
- Hartley, H. O. (1940): Testing the homogeneity of a set of variances
Biometrika, **36**, 317–346
- Johnson, N. L. and S. Kotz (1970): Distributions in statistics: Continuous univariate distributions 1
John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore
- Kendall, M. K., A. Stuart and J. K. Ord (1987): Kendall's advanced theory of statistics, Vol. I: Distribution theory
Charles Griffin & Company Limited, London
- Kendall, M. K. and A. Stuart (1979): The advanced theory of statistics, Vol. II: Inference and relationship
Charles Griffin & Company Limited, London
- Kullback, S. (1934): An application of characteristic functions to the distribution problem of statistics
The Annals of Mathematical Statistics, **5**, 263–307
- Lehmann, E. L. (1986): Testing statistical hypotheses
John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore
- Lentner, M. M. and R. J. Buehler (1963): Some inferences about gamma parameters with an application to a reliability problem
Journal of the American Statistical Society, **58**, 670–677
- Linhart, H. (1965): Approximate confidence limits for the coefficient of variation of gamma distributions
Biometrics, **21**, 733–737
- Nair, U. S. (1938): The application of the moment function in the study of distribution laws in statistics
Biometrika, **30**, 274–294

- Nayer, P. P. N. (1936):
Statistical Research Memoirs, **1**, 38–51
- Moran, P. (1951): The random division of an interval - part II
Journal of the Royal Statistical Society, Ser. B, **13**, 147–150
- Schremmer, D. (1986): Ergebnisse zum Ornstein-Uhlenbeck-Prozeß und zum Test
von Anderson und Darling
Forschungsberichte aus dem Institut für Statistik und Wissenschaftstheorie
der Universität München, Serie Sp. Nr. 4
- Shorack, G. R. (1972): The best test of exponentiality against gamma alternatives
Journal of the American Statistical Society, **67**, 213–214
- Spanier, J. and K. B. Oldham (1987): An atlas of functions
Hemisphere Publishing Corporation, Washington, New York, London

Lebenslauf

Name	Rainer Vollmerhaus, geboren am 15.9.1959 in Gießen	
Eltern	Bernd Vollmerhaus Christel Vollmerhaus, geb. Schulte	
Schulbesuch	April 1966:	Volksschule Gießen
	November 1967:	Volksschule Neufahrn
	September 1970:	Josef-Hofmiller-Gymnasium Freising
	Mai 1979:	Ablegung der Reifeprüfung
Studium	November 1979:	Diplom-Statistik an der Ludwig-Maximilians-Universität München
	Juni 1985:	Ablegung der Diplomprüfung
Beruf	Seit dem 1. Juli 1985 wissenschaftliche Hilfskraft am Seminar für Spezialgebiete der Statistik an der Ludwig-Maximilians-Universität München	

